Customized FORM PTO-1390 U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

ATTORNEY DOCKET NO. P06780US0/TPS

U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 09/581422

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/EP98/08557	14 DECEMBER 1998	15 DECEMBER 1997

TITLE OF INVENTION: A NEW GENE CALLED OLIGOPHRENIN 1, ITS EXPRESSION PRODUCT, AND . . .

APPLICANT(S) FOR DO/EO/US: CHELLY, Jamel et al.

Applicant herewith submits to the US Designated/Elected Office (DO/EO/US) the following items and other information:

☒ 1. This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.

☐ 2. This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 USC 371.

☒ 3. This express request to begin national examination procedures (35 USC 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 USC 371(b) and PCT Art. 22 and 39(1).

☒ 4. A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.

☒ 5. A **copy** of the International Application as filed (35 U.S.C. 371 (c)(2))

- ☐ a. is transmitted herewith (required only if not transmitted by the International Bureau).
- ☒ b. has been transmitted by the International Bureau.
- ☐ c. is not required, as the application was filed in the United States Receiving Office (RO/US).

☐ 6. A **translation** of the International Application into English (35 U.S.C. 371(c)(2)).

☒ 7. Amendments to the claims of the International Appln. under PCT Article 19 (35 USC 371 (c)(3))

- ☐ a. are transmitted herewith (required only if not transmitted by the International Bureau).
- ☐ b. have been transmitted by the International Bureau.
- ☐ c. have not been made; however, the time limit for making such amendments had NOT expired.
- ☒ d. have not been made and will not be made.

☐ 8. A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).

☐ 9. An **oath** or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).

☐ 10. A translation of the annexes to the Int'l Prelim. Exam. Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

☐ 11. An **Information Disclosure Statement** under 37 C.F.R. 1.97 and 1.98.

☐ 12. An **Assignment** document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included

☐ 13. A **First preliminary amendment**.

☐ A Second or subsequent preliminary amendment.

☐ 14. A substitute specification.

☐ 15. A change of power of attorney and/or address letter.

☐ 16. Other items or information:

- ☐ **Small** Entity Statement
- ☐

☐ A copy of the Notification of Missing Requirements under 35 U.S.C. 371.

☐ In the event that a petition for extension of time is required to be submitted herewith, and in the event that a separate pet does not accompany this response, applicant hereby petitions under 37 CFR 1.136(a) for an extension of time of as m months as are required to render this submission timely. Any fee is authorized in 17(c).

Date: 13 June 20

LARSON & TAYLOR, PLC • 1199 North Fairfax St. • Suite 900 • Alexandria, VA 2

4.0 Rec'd PCT/PTO 13 JUN 2000

U.S. APPLICATION NO. *(If known)* 09/581422	INTERNATIONAL APPLICATION NO. PCT/EP98/08557	ATTORNEY DOCKET NO P06780US0/TPS

		CALCULATIONS *PTO USE ONLY*
☒ 17. **The following fees are submitted:**		
☒ **Basic National Fee** (37 CFR 1.492 (a) (1)-(5):		
☐ Neither Int'l Prelim. Exam. fee nor Int'l Search fee paid to USPTO	$ 970	
☒ Search Report has been prepared by the EPO or JPO	$ 840	
☐ No Int'l Prelim. Ex. fee paid to USPTO but Int'l Search fee paid to USPTO	$ 690	
☐ International preliminary examination fee paid to USPTPO	$ 670	
☐ Int.'l Prelim. Ex. fee paid to USPTO & all claims satisfied PCT Art. 33(1)-(4)	$ 96	
ENTER APPROPRIATE BASIC FEE AMOUNT =		$ 840.00
☐ Surcharge of **$130** for furnishing the oath or declaration later than ☐ 20 mos. ☐ 30 mos. + from the earliest claimed priority date (37 CFR 1.492(e)).		$

CLAIMS	*NUMBER FILED*	*NUMBER EXTRA*	*RATE*	
Total Claims	DEFER - 20 =		X **$18** =	$ DEFER
Independent Claims	DEFER - 03 =		X **$78** =	$ DEFER
☐ Multiple Dependent Claim(s) (if applicable)			+ **$260** =	$ DEFER
TOTAL OF ABOVE CALCULATIONS =				$ 840.00
☐ Reduction of ½ for filing by **small** entity, if applicable. A Small Entity Statement must also be filed.				$
SUBTOTAL =				$ 840.00
☐ Processing fee of **$130** for furnishing the **English translation** later than ☐ 20 mos. ☐ 30 mos. + from the earliest claimed priority date (37 CFR 1.492(f)).				$ DEFER
TOTAL NATIONAL FEE =				$ 840.00
☐ Fee for recording the enclosed **assignment**, accompanied by a cover sheet - **$40** per property				$
TOTAL FEES ENCLOSED =				$ 840.00
Amount to be				*Refunded* $
				Charged $

☒ a. A **check** in the amount of $ 840.00 to cover the above fees is enclosed.

☐ b. Please charge my Deposit Account No. 12-0555 in the amount of $ to cover the above fees.

☒ c. The Commissioner is hereby authorized to charge any additional fees required or credit overpayment to Deposit Account No. 12-0555.

Note: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

THOMAS P. SARRO

At the address (below) of **CUSTOMER NO. 000881.**

LARSON & TAYLOR, PLC
1199 NORTH FAIRFAX ST.
SUITE 900
ALEXANDRIA, VA 22314

SIGNATURE: Douglas E. Jackson

NAME: DOUGLAS E. JACKSON

REG. NO.: 28518

PHONE NO.: 703-739-4900

Date: 13 JUNE 2000

09581422 061300

SEQUENCE [L1]LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Institut National de la Santé et de la Recherche Médicale
(B) STREET: 101 rue de Tolbiac
(C) CITY: Paris
(E) COUNTRY: FRANCE
(F) POSTAL CODE (ZIP): 75013

(ii) TITLE OF INVENTION: oligophrenin 1 gene and protein

(iii) NUMBER OF SEQUENCES: 27

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1650 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TTTTTCAACC CATGGTGCAT TTAAAGCTCA TTTTGGACAT TTTCCCATAA ATTAACCTTA
60

AAGATAAAAA GAGTAAGAAA CAAAACTTTC CCCTGAGATG TGGCTACTTA TTTATTTTCA
120

GAGGGCGTTT TCTCATTGCC ATTCTGCTAT ATAGTGTAGT GGTCAAGAGC ACTACCTCTA
180

GAGCCAGCCA GGCTGGGCTC AAGTTCAAGT GCTGCCATTT AACTAGCTGT TTGTCCTTCG
240

GCAAGTCACT TAAACTCTCT TTGACCCAGC TTCTCCATCT TTAAATGGGT ATAATAATAA
300

AACCATCCTC ATAGGGTTGT TTTGAAGATT AGTGAGATGG GCGATAGGTT GTGTGGTGGG
360

TAGAATAATG TTTCCTCCTT CACAGATGTC CATGTCCTGT CCTGAAACCT GTGGCTACGT
420

TATCTTATGT GGCAAAATGA AATTTGCAGA TGTGATTAAG GATGATGTGA TGGGGGAGAT
480

TATCCTGTGG ACCCAGTGTA ATCATAAGGG TCCTTAAGGG GAGGCATGAG AATCACAGTG
540

ATGTAGCATG AGAAAGACTT GACTGGCCAG TGCTAGCTTT AAAGACGGAG GAAGGGAGCT
600

ATGAGCCAAA GGATGTGAGC AGCTTGAGCC GGAAAAGGTA AGGAACCAGA TTCTACCCCT
660

AGACTATCTG AAAATGAACA CAACCCTTCT AACATCTTGA TTTGATCCCA GTGAGGCCCA
720

TTTCTGATTT CTGCCTTCTG GAACTGTAGA TTATAACTTT ATGCTGATTT AGGACATCAA
780

GTTTGTGGTA ATTTGTCACA GCAGCAGCAA CAGGACCTAA TTCAGGTGGG TCAGCCACTA
840

TTATTCTATT TACAACCCCC TTCCGTTCTG GCTTTCTTCA CAGGGCAATC TAATGTAAAC
900

ACTTCTGCAG AATTAAAGGG ATCTCTTTGC CCACGCACAA ACTATTTTTT AGTTTTTCGA
960

GCCTCTACCT ACCCTTAGTC TCAAGACAGC TGCCTTTAAG GTAAGCTGCA AGACAATTTA
1020

TAACATCCTA TTCTTAGACT TTCCACTTAC ACATGGAAAA TCATATGCAT TATCGGCGTC
1080

TGCCACAAAA GCCATGCTCT GAGGTCCTAT TCAGGAAAAT ATTAACTCTG AATTATGGCC
1140

TCCCTTTCTT TGGCAGGGGG CACTGAGCTT CACGGAGACC CTAGAGTAGG AATCAACCCT
1200

TCAGGTGGAA GGCCCAAGGG AATTAGAGCA CTTAGAGCAG GAGACATGGG AAAAGAAAGA
1260

GGAAAAGAAG TTGGTAGGTC CGAAGGAGGC TGGAAAGAGG GGAAGGCAAA AGAATACACA
1320

ACCCAGCCGA ATCATGGGAA GTGAGAGATG GCTTCCTGCA ATCTAGTCTT GCGGGTTCTT
1380

GCTCTAGTTT CGGAAACTTC CCCCGCAGAG TCCAGCGCCG CGCACCTGCC CCTTGGGGCC
1440

GGTCCCTGCG CAGGAGCCTG CGGCAGGCTG GGAATGCCGT TCTGGGAGAG CAGCCTCCCA
1500

GGCGGGGTGG TGGGGAGACT GCAGGGAGGA GTTTTGGGGA GTGCAAACGG AAAGACTACA
1560

TTTCCCAGGC CGCCACGCTT TCCAGCTGGA GTCCTAGGGC GCTGACTGCT CCCCAGTTTC
1620

CGTAGGGAAG CGCTGGGCTA CCGCGGCTAT
1650

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1079 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 1
(B) LOCATION: 1..634

(ix) FEATURE:
(A) NAME/KEY: exon 2
(B) LOCATION: 778..935

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 727..746

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 958..977

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TGTGCGTCGC GCTCTCGCCC TCCTCTTCCC GCTGCAGTGT CTATGGAGCG AGGCTACGTT
60

TCATTGCCGC CCTGGCTTAA CCCTTCCGGC GCCTAAAAGG ACGGCCGGCC GGCCGGTCCC
120

TTGCACCAGG AAGAAGTCTT AGCAGCCAGC GGGCCCTGGT CAGGAAACTC TAAGGTACAA
180

GGAAAACAGT TGAGGAAGGA GCCAGAGCGC TCCGGTTTGG TCCTCGGGCT TCGCTGGGGC
240

GGGGCGCAGG CGTTGGCTTT AAGAAAGGGG AGGGGACAGT GCAATCCGGG TTGCCCGCGG
300

ATTCGGCCAA GGAATCTTCC GCTCGCTCCG GAGCGAGGAG CCTGTAAAGA GGCTGTTCCC
360

AGCTCCAGCT CTAACCTCGC CTACACCTTG GGCGGGCCCA ATGTCACGTT TGCAATTGCT
420

CAGGAAGGAT CCGGCCCGTC TCCGGAGGCA AGTCGGGCTG CGGTTTTTGC TGCTTATCTG
480

GGAAGGCGAT GCCTAAGGGA CATGCTGCTT GCTAGGCAGC ACCCTGCCGG GATCCGACTG
540

CGATAGTTAG CTCTCCCTGG CCCTGAAGCC ATCGCCGGGG CGCCTGTTCT CTGTCCGGAC
600

CAGCCAGCGC TCCTCAGGAG TCTCACTGAA ACAGGTACCT GTCCTCCAAA GGGACGGAGG
660

CTATGAGCTT CCTTAAGCGG GTCGCGCGCT CAGTCCGTCC CCTCTACTTC CTCTACTGTG
720

CCATTGATGC TCTCGGTCTT TGTGTCTTTC CCCTTTCCCC CTACTCCCGG CCATCAGAAC
780

CATGGGTCAT CCCCCGCTGG AGTTCAGCGA CTGCTACCTG GACAGCCCCG ATTTCCGCGA
840

GAGGCTCAAG TGTTATGAGC AGGAACTGGA GAGGACCAAC AAATTCATCA AAGACGTAAT
900

CAAAGACGGC AACGCGCTTA TCAGCGCTAT GAGAAGTAAG TGCAAGGCTT CGATGAGCTG
960

TTTCTCTGAN CTGGTGTGTC TGGCCTTTAA GCCTTTCCAC ACCACCAGGG GAAGGGAGAT
1020

TGCAGGGTGA CTCCCAGCCC ANATCTCTGA GGCAAATGGG TTTCCCACAC TTGGGGAGT
1079

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 900 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 3
(B) LOCATION: 403..498

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 375..394

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 504..523

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CTTTGAACTC CACACGTATA AAATGGAAAG CATCTAGTGT ATTGCCCATA ATAGGGGTTT
60

AGGAAATGTG TAAGGAGACC CTCTTAAGAG TATATCTAAG TAGTTTCATG TCTTCCATGC
120

TTTGAGTGGA ACAGGTCAGA GAGGAGAGGT GTTGAGGATA AAAACATGTC CCGTAACTTT
180

TAAGGACTTT ACTGAGATGC CCCATCCTTT CTTCTTGTGA TTTTAGTTGT TCAAACTTTC
240

TTCTACTGCA TATCTAATCT TTTGTTTTAT TTCATTAAAT GCTAGTTGCA ACCTGCTCAG
300

AGCACTTACC TTTGGTTTTC TTTTATCTGC TCTTATAGAG ATGAGGAAAT AGATCAAAGC
360

ATAGCATTGT TGAACATTTC TCTTGTTACC CTTCCTTTGC AGATTATTCT TCTGCTGTTC
420

AGAAATTTTC CCAGACGCTG CAGTCATTTC AGTTTGATTT CATTGGAGAC ACTCTGACTG
480

ATGATGAAAT TAACATCGGT AAGTCTTCAG CTACATGTGG TCATATACCT GTTGAGGCAG
540

CCCTGAGACC ATGTAGTCTT TTTGATTTGT GGATACAGAG CACTTGGACA TCTTCATCCA
600

CTGTGGTCCA ATGCCAAGGC CCTGGGAGGT TGATTAGGAA GGATCAGGAA ACTTTCCCTG
660

CCAGTCCCAT TTCCTCCTCA CACGACAGCA ATCAAAAGAT ACCCTTAAAC TTCTACTGAG
720

ATTTTTGACT CAGACAGTCT GCAAGCGACC TTTTCTTTAA AGCATAGTTA TTTTCCTAAA
780

GGATATATTA AAAGGGGGAC TTTTTTGGTT TATTTCCCAA AATGGTTGGA GTTAGATTCT
840

TCTAAGGAAT CAAATTTCCC TAGAAAGTGT TAAATTAGCA TTTGTGTGTC TACAACTTAT
900

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 960 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 4
(B) LOCATION: 483..544

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 418..437

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 551..570

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TAGAAGAGTA GAAAGTTTGG GAGGGTGGAA AGGGTAATGT TTATTATTTA CTTTGTGTTG
60

ACACTACATT CACCATCCAA TTTAGTTTTT ACAGCAATAG TCTCTGGTTG GTCAGTAGTG
120

GAACAGGGAT CCAAACTCGT TTCAGGGTTC TATAGCTGCC AATTAATTAC ACAGCAAACC
180

TCTTGCCCCT TCCCTAATCT TTTAGCTTGT CTATGAAATA AGAGAAATTC TAGCTACTCC
240

TGAGTGGCTG TAAGGATTAA ATAAAATATT AAAGTGTATG GGGATTGATA AAAGAGGAAA
300

GAAAAGAAAA GAAACATTCA ACAGGTGCTG AACACCTGCT TTTGTCCTCC GATTTGACAC
360

CTTCCTCTAG TGGCCATGTG GGCATAGGGC ACTGGTCCCT ACTTCCTGTT GCACAGATCT
420

CTATCCATTT GTCTGTCAAG CACCATGATT AATTTGTTTT ACATTTGATT CTCTCCTTCC
480

AGCTGAATCC TTCAAGGAAT TTGCTGAATT GCTCAACGAG GTAGAAAATG AGAGGATGAT
540

GATGGTAAGT CACTAACGCT GTCACTGAAG CTGAGTTCAT GGGTGATATA GGGGATTTTT
600

CCTTTCCTTT ATGCTTGGAT TGATCCTATA CTATTTTGAT TTCTGTCAGA TAGCTTCTTG
660

GTGCTATAAA AATAGTTAGG TAATAGATCT GGTTATTATG TCTCAAGCTT CCACCCTGAG
720

AGTTTGGCAT TAGATAGAGG GAAATTAACG TGCAAATCCC ATCTGTGTTC ATTTCAGTGA
780

AAAATAATTT CAGTGGATAT TAAACTGGGC CTTTGAACAT GTTGACAGAA ATTGAGGTCT
840

TTAGTGTTTT TAGCCAAATT ATCCATTTGT TAATCTTTAA TTTGTGGAGT AGTTTTACTT
900

TTATAGAGAA AATCAGTAGA AAATAAAGAT AGAACTCATA TACCACCTTT CTCTCTCCCA
960

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 960 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 5
(B) LOCATION: 451..522

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 423..445

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 553..574

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TCTGGTTTTA TGGCACAGGC AGGACTGTAA ATGTTTTATC CTAATGCTAT CACTGCACAC
60

TCATTTGCCT CTGATGAAAT GATTTGCACT TGCTGCAATT GTCCTTTTCT TGTATTTGCC
120

ATTCTCTTTC TCTTCTCTTT TTCACCTGTC CTTCAGGTCT CCTGTGTCCT CTGCTGGACA
180

TCACTTCAGA TATTTATCGA TTAAAAACTC AGGTCAGACT ATCATTAAAG TTACAGAGAA
240

ATGCCCCTCT TATTCTTTCT CCCATTTCTT CTCAATGCAT TTGATTTTTC AGAAACAATA
300

TAGAAACAAA CAGTAACAAA ACCCAACAAA TCAGCAAACC ATTTAACATT TTGCAGGTTG
360

GTATATAAAT GAAAATGTAG TAACAAGGAA TCTTGTATCT GAACCTTGTT AACCTAGAAA
420

TTGTTTTGTT TGTTTTTCCT TTTTGTCTAG GTACACAATG CTAGTGATTT GCTGATTAAA
480

CCCTTGGAAA ATTTCCGGAA GGAACAAATA GGCTTCACCA AGGTACATTT TCTGTATATG
540

CATAAGATTT TTTAAAATAG CAATCGAATA GTTGTATGGG CTACTATTCT TCACTTTACA
600

AAGATATGCA CCAATCTGCT GGTGCTTTGC TCTTGGCCTA GTCAGCCTCC TAAACTGTGC
660

AAAATAAATG TTTGTTGTTT ATGTCACCCT GTCTATGGCA TTCTGTTATA GTAGCCTCAG
720

CTAACATGAC AAAGGGGGTG GGGAGGTGGG TGATTAGTTT CTATGAGAAA ATGATCACGA
780

AAGAGAGTAA GAAAATCTAG AATTGGCCTC TGACTTTGTG GCCAACAGGC TCTGTATCTG
840

TGCATAAGTT TCTTCTTCTT TTGGTGTTTT TGTTGTTTTG TCTGGAAAAC TAGCTAGCTA
900

TCATGTATCA ACTGCCTGCT ATATTGAGCA CTAGGCTAGG TGCTTTACGT TCATTCTTTA
960

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1020 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 6
(B) LOCATION: 416..517

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 388..407

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 540..559

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CATTGTAGAT TAACCTTTTC ATGACAAATA TTTACTTCCA TTTGTGTTAT GTCTTCTTCC
60

TTTGTAAGAA AATTTGAATG AGTGGTTAAT CTATGTGAAA AATATTGGAG GGAAGAAAAT
120

ATATCTACTG CACAGGCCCT TTTAAGGTAT CATTCTCTAA GGAGCAGCTT CCATAGCTTT
180

CAGCTGTAAA AATAGGGACT GCCATTTCTG CAGGCAGAAT GGTTTGGGGT TATATTTCAG
240

GAAGCTGAAA CTGCTGAGAC CAATACAAAG TTAATTCTCC GTTGCTTTTT TTTCCTTCCA
300

GGAGCTAGAG GGATTGACCA CCTGAAACCT GACACTATCT CCTTCATTCT CCTTTCTAAG
360

CAGACAGAGT TACAACTACA GACATTTAAT CTTTGCCCTT CTTTCCCACC TTTAGGAGCG
420

GAAAAAGAAA TTTGAAAAGG ATGGTGAGAG GTTTTATTCT TTACTGGATC GGCACTTACA
480

CCTGTCTTCA AAAAAGAAAG AATCTCAGTT ACAAGAGGTA TGTTCACAAA GCCTGCCCCT
540

GCCTTCCATT GCTAGCTATG CCTTAGAAAC AGTGTGAATT TTGTACTGCA AGGCTTTTCC
600

CATACCCCGT CTCAGCAGGG AACCTCATGT GATAGTAGCA CTTGTAGTCA AAACTGTGGC
660

CTGAGACTCA GAAGCCCTGA ATACTAAGCC AGCTCTTCCA CTAACTCAGG GTGTGACCTT
720

GGATAGGATA CTTCTTTCTT TGCCTCATTC CATTATCTGT AAGAAGAGGA GTCGAGAGTC
780

CCTTTCAATT GCAAGTCCAA AATCCATGCG AGGATAAAGT TAAACTAGTG TTGTATTTGG
840

TAGAAATCAG GAACAGATCT TCTACTTTTT TCCCTGAGAG ATTCCACAAC CTTTTTTTTT
900

TTTTTTTTTT TTGGTGAGGG GTTAGGGGAA TGTTCTGTTT TGGGACTACT GGTTACCTGG
960

GACTTGCAGT GCCTTCAGTT CAAATAAGCT ACCATTCGGT GAGACCCTAC CACATGCCAT
1020

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1020 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 7
(B) LOCATION: 464..574

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 436..458

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 584..603

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CATTTCTTGT TAAATAAGAC TTGAGACTTG CAGTTCTTTC TGTTGTAATC CCCTCAGTAT
60

TGGCCATAGT AGCATGCTTC TGTTCTTGGG CTTTGGCCCT TGTTAATTTC TTATTCTCTG
120

TATTCATGTC TCTGTAGTGT TTGAGGGAGC AGATTTTCCT CTGACTTCAG TTCTCTGTTG
180

AATATAATAA AATGTGTGGA TTTTACACTT GTTCAGCTTT TTTCTTTTTG TCAGGGTGGA
240

TGTAATGACT TCAAGCTTTT TATGTGTTGG ACCAGAAACT GGAATCCTGT ATCAGTCACT
300

TTTTTATCTC TTGGCTTTAG GTTCTCTATT TCTAAAATAA AAGGTAACAT CACTAGTTGG
360

TGGCTAAATG CTCATCTAGC TCTAGCATTC TGTTCAATCC TTAAGTACTG ACTTTCATGA
420

ATGAATATCC CAATATGTAA TGTTTGTTTT CCTTCTTCTA CAGGCAGACC TACAGGTGGA
480

CAAGGAGAGG CACAATTTTT TCGAGTCCTC TCTTGATTAT GTTTATCAAA TCCAGGAAGT
540

TCAGGAGTCC AAGAAGTTCA ATATTGTGGA GCCTGTAAGT TTTCTCTGTT GATGAATGGT
600

CTAAAAATAT TTATCAAATG CCTGGTAAAT GTACAAACTT TGATCATAAA AACAGAATCC
660

CAAAGAACGT GAAGGATAAT CAGTGGAAAT GTCTAGAACG TAGTCTAACT TACGTTTTTC
720

TTGTTCTTTA GTATACTTTT TTTCTCTCTT TCCCCCCAAC CCTTTTCTCT CTTTTTCTAT
780

GGCTATTTTC TTTCTCTTTC CTTTTTTTCC TTTCTCTTTC CTCCCCCCTT CCGGAATTTC
840

TCTACAACTC TTCCTGTTAC TTTGTTTCCC TCCCTTCCTT CTTCTTGCGC TTTCAGAAGC
900

TTCTGAAAGT CTTTCTCATA TACCAGATAC TATGCTAGGG ACTCATGTCT CTTGCAGTCA
960

GTGACTTCTC ATTCTACTAC TTTTACTTCT GTCTTGTGTA GTTTTTTCCC ATTTCACGTT
1020

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 479 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 8
(B) LOCATION: 244..348

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 219..239

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 363..381

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCGCCACCAC GGCCAGCTAA TTTTTGTATT TTTAGTAGAG ATGGGATTTC ACCATGTTGG
60

TCAGGCTGGT CTCAAACTCC TGACCTCACG ATCTGCCCGC CTCGGCCTCC CAAAGTGCAG
120

GGATTACAGG CATAAGCCAC CACACCCAGC CACACCATGG AGTTTTTTGG TGAGTTCATG
180

TTTCTTTTAT TTAGTTTATT AGAAGATGCT GGTGATAAAG TTATTTTTAC ATGTTTTCCC
240

TAGGTCTTGG CCTTTCTTCA TAGTCTGTTC ATTTCTAACA GCCTGACTGT GGAGCTCACA
300

CAGGATTTCC TCCCATACAA ACAACAGCTC CAACTCAGTT TACAGAATGT GAGTTTGCAT
360

GTGGATTTTT CTCACCGGTC TTTCCATTCC GATTGAATTT CAGCCCTAGC GACCTTGATT
420

CTTGGAATTC TAGGTTACTG CATCCTAGCC AATTTGTTAG AATATACTGG TGTGGATCA
479

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 600 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 9
(B) LOCATION: 134..263

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 108..128

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 336..355

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ACCATTTCTT TTTTTGCTTT GTGGTTGCAC ATGCTGTAAG CAGGGAAAAC TTTGTACTGA
60

GTCTCTGACC AAGAAATACT TTTTCATGAT AATGATGATG ATAATAATGA TTTTCATGAT
120

GATGTCTTGA CAGACAAGAA ATCATTTCTC CAGTACCCGG GAAGAGATGG AAGAACTTAA
180

GAAAAGGATG AAAGAAGCTC CCCAGACATG CAAACTTCCA GGACAGCCAA CTATTGAAGG
240

CTATCTCTAT ACACAAGAGA AATGTGTGTG GGGACATAGG GGTATCCATT GGGTTTCAAT
300

AAGCCAGGAA GTACTGCCAC TTGTCGGCTG TGAATTTTGG GCACCCTTAC TGTTCATAGA
360

CCCCTGATAG CTAAAATTCC CTTGGAACGC AGGCAGGGAA TACTGAAAAC AAAAAAAAAA
420

AAAGGAGAAA CTGAGAGGAA GTTAAAGATT TGTCTTACAA AGGCTGTGTA GTGATAAGAC
480

CTAAGGTTTT CTCTGAGATT CAAAATGGGT ATTATTTGTT CTTTAATCCT TCTGATTATT
540

CTTTTGGAAA AAAGGGAAGT AGAGGAAAGG AAGTAGAAAA ATAATATTTC TTATACTTAT
600

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1020 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 10
(B) LOCATION: 383..483

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 361..380

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 492..511

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGAGAGATGT GTAGAGTCAT GTACAGGTTA GAAGGGGCAT AACCCAATGA CCAATGACAT
60

GGAACAGTTA TAGAGACAGC AAAGTAAATG ACAATAGTCT CCAGCTTCTG GGATATATGG
120

TGATACAATT TATAGGGAAC ATGGCAGGGA GAGTAGGTTT TGAGCAAGCC AGAAGCAGCT
180

GTGGGAAGCA GTTGGTGAAG GTTAGAATCT CCCTGTGTGC AGTAGGTAGG TGGCTATGGA
240

AGGAGGGCGG TCAGGGCAAG GGCAGGGCTG GATCTGAAGT TTGACTCTGA AGAGCAATGT
300

GTAAATAGCT TCCATCTTAG GGTTGACTTC CTATACAGCT AAAATAGTTA TTCTGTCTGC
360

TCACTTTTAC TTGTCCTTGT AGGGGCTTTA GGAATATCCT GGGTGAAATA CTATTGCCAG
420

TATGAGAAAG AGACCAAAAC ACTGACCATG ACGCCTATGG AGCAGAAGCC AGGTGCTAAG
480

CAGGTCAGTT CTTGTTTGCA CCATATTTTT GGAAATGGAT CTATGACTGT TTCTCAGAAA
540

AGAATATATG TTGACCTAGT ATCAAATCAT CAAGTTCATC ACTGTTACGT GAGGCCATGA
600

CTTTATATGT ACACCTTGGC CTAAGTTTGA GTCAGATAGC ACTGAGTTGA GTGAAAAATT
660

TCTCTGTTGA TTAGAGCAAG CCTTTTGAAA GTGCCGGTAG TCTTTCAAAC CAGTTATTTT
720

TACAAGTGCC AGTCACATTG TACAGTCAAC TATGTAAAAA TATGGATGAA TTACTTTTAA
780

GAATGCTCTA CTCTTGGATT CTTTAAAATA GCAAGTTTTA AAAATATGAA TTGAATTCCA
840

AAATTCCTTT TTTACAGGAG TGTGTTTATG GCCCACAGTT GGAATAACCG ATACTCACAT
900

TCTATGTACT ACTCAAATAT CTTTAAGCAG TTAATCTCTC TTTTTCTGCC CTCCAAACCT
960

TCTCTCACTC CTGAAAATGA CAAGATAAAT TTAACACACT GAAAAAAATA GTTTACTTAC
1020

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 294 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 11
(B) LOCATION: 107..198

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 81..100

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 223..242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CCCTAGACCA GACGCCTTCC AAGTGACACT GAAGACACTG TAGGAAGCAG AAACAGCTTC
60

TGTGTCTTTT TTTAAAAATT GCCTGTTAAT TCTTGTTTCA TATCAGGGGC CCTTGGACTT
120

AACACTGAAG TACTGTGTGA GAAGGAAGAC GGAGTCTATC GACAAGAGGT TCTGTTTTGA
180

CATAGAAACT AATGAAAGGT AAGCTGTGCC GCTGTGAATT GGCAATGTCC CCACGTGCCA
240

GATGCTTAGC CTGGGTATGT CTTTTATTTT CCTCCGTCAT CCCACGTTGA TGAC
294

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 494 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 12
(B) LOCATION: 211..289

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 188..207

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 300..319

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CCACCCCCGC CTCCATGGTT CCCAAAGTTA CTGTTCTGTA AATTGCCTAT TGTTTTTCTT
60

GAGTAAAAGA GATTTTGAAA AATTAGTATT CTGTAAGCCT ATGCTTGTTC AGTGTGACAA
120

ACTCCAGGGC AGAGAGGGAC CCTAGATCAC AAGACTCCAT TCTCTCAGTT GAATTTTCTG
180

CTTTATACTT ACCATTTTTT TCCCCCTCAG GCCAGGAACC ATCACTCTGC AGGCCCTTTC
240

AGAAGCTAAC AGAAGGCTAT GGATGGAAGC CATGGATGGG AAAGAACCTG TAAGTTACCT
300

GACACTGGGG CAAACCTCCC CAGCATATGC CAGTGTATGA GTGCCCTCTA GTGGTATCAG
360

TGGGTCTCAN ACAATTAAAT GGTAATGGAT TGTTTAGTCT CAGTTTTAGA GCTGTAAGGA
420

ATTGTTTCCA CATCTCTTAG CAGGTAAGGC AACTGGAGTT CCAGAAAGGT TGAGGGACTT
480

TTCTGAGACC ACCC
494

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 378 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 13
(B) LOCATION: 212..245

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 166..189

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 259..278

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGCATGAGCC ACCACGCCTG GCCTGTTCAA GTATTTTCTA GCAATCTTGG CAAAGCAATT
60

ATGTTTAGCC CACTTGGCTA TCTTTTTAAC ATCCTGGAGT TTCTAATCAT TTTTAATGCC
120

TATCTGGGGA AAGATATTTA ATATTATGTT CTCTGTTTTC CTATATTGAT TGACAATAGC
180

CATGGATCTT TCTGTTTATC TTCTTTTGTA GATCTACCAC AGCCCTATAA CAAAACAGCA
240

AGAAAGTGAG TCACTTAAGT TTTTGGTCTA CTAGCATTAT AAACTGCCAG CTGTCCGATT
300

CATAGTAAAT ACCATCATTA ATGATGTGTA CTACTAACGC AAGTCTGAAT ATGGATGCCT
360

TTGTGTGAAA TAAAATTC
378

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 393 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 14
(B) LOCATION: 172..234

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 133..152

(ix) FEATURE:

(A) NAME/KEY: oligonucleotide
(B) LOCATION: 250..269

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

AAAATGTTAA ACCTCCCCTG AAAAAATGAC TCTTTCCATT TAAGGGTGAC TAGAAATGAG
60

CAACTAAAAA CCCTTAGCTC TCTCAATGCA GTCCCTTTGC ATGGTCATTA AATGTTTAAT
120

AGGTGACACC TGTTGCAGCA GGATCTAACT CTTTTCCTTT GCTTGAAACA GTGGAGCTAA
180

ATGAAGTGGG CTTCAAGTTT GTCAGGAAGT GCATCAATAT TATTGAGACC AAAGGTAAGA
240

TCTGAACCAT AGTCTTGGCA TTGTCTGAAT CTCGTCACTC TGATTTTATC CTGGGCAATT
300

TCTCTGAAGT AGCGTTTTAG GAATGAAGAC TGTTTATAAA GCTTGTGTAG TAGATGCAAG
360

CTAGAAAATT TCAGAAAATT CTAAACTAGT GGT
393

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 436 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 15
(B) LOCATION: 207..281

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 151..170

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 293..315

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

AGATTAATGA GGGTTTGGTA CACTCCAAAT GAAAGGATGG ATAATTTGGA GAGATGCTGT
60

AGGACTATTC CCCTGTTACA GGGAAGGCTG GAGAACTTGG AGTATGTAGT GTGACCCCTT
120

CCTATCTGAA TTGACTCTAG TGTACCAAGG GGAGATGACA ACTTTAGCTA TACAAGTGAA
180

ATTAACCTGA TTTTTTCCTC CACTAGGGAT CAAGACAGAA GGGTTGTACC GCACTGTGGG
240

CAGCAATATT CAGGTTCAGA AGCTGCTGAA TGCCTTTTTT GGTAACAATT TCACTTTGAT
300

AATTCTTATT GGGAGTACTT TATGTGTTAC AAAGAAATGT GACTGGAAGA GAAAGGAGAC
360

ACTGCTAAAA TGTGGTAGAA TAGTTGAAAA AAGTATTTTC TAAAGTAAAA CATACACATA
420

CTTGCCCACC CTGGGC
436

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 547 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 16
(B) LOCATION: 270..354

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 221..244

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 363..382

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

ACCCCAGTNA TNTGATGAAT CTAAGAAGAG TNGATNTTGT TTGTTCAGTT TTTTCTTGTT
60

GTGTGAATTG GATAGATTAC TTTNTTATTT CTTATATGGC AGACCAGAAT GCAGTCATGT
120

TTTTGAAATA TCAAAGATTT GCTTCTTCTA AAGTTTTGAT NTCTTAAAAA CTACTTAGGG
180

TNATATACTT TGTTTTTCTT TTAAAAGAGG GAAAATGTAA GATTTTTTTG ATGATTAACT
240

TTTGTTTTTT GTTTACTTTT CTCAAATAGA TCCTAAATGC CCAGGAGATG TTGATTTTCA
300

TAATAGTGAC TGGGACATTA AGACAATCAC CAGCTCCTTG AAATTCTACC TCAGGTATGC
360

CTGATTTGAA TTGGGAGTTT GCTTTTCATA GCTGGTGAAA TTTCTCTGGG TGTTGAGCGG
420

AGTTAACGTG GTCTCAGTTC CAGGAGTTTG GATACAATTG CTTAANAAAA AACATGTGAA
480

GAGGATTTCT GGCCANGAAT GTGCAAANAC TGTTTTTTAA ATCTGAGAGT TTAAGCAAGA
540

GAAGCAT
547

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 601 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 17
(B) LOCATION: 355..413

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 305..324

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 438..457

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GGACAATTGC AAAAGCACTT CGGAAATTCT AAGGATCTAT CAAATCGTAA GGGATTCATG
60

GTAGCATTCA GCATGGGTCC CCTCTGGAAT TTTGCAGGAC TGGTTTGTGC CTCTTTTTAC
120

TTNTGGGAGC TAGTTGGAGA CCTTGCTAGA GGGCTCAGCC CATGCTTTTG CAGGTCTTTT
180

GTTGAATTAC TAGCAACTTG GATTCCCTGA CGAAGCTTCA GGTGAAGAGA AAAATGTATA
240

TAATCCCACT AAGCTGTAGG GCTCAGGAAC TTCAGCCTTG CTGTCCCCAG AACTAAGAAT
300

CCAATACCCA GCTGCTTTNT TCCCAAAGCA ACTGACAATT TTCATTCATT TCAGGAATCT
360

TTCTGAACCT GTCATGACCT ATAGACTTCA CAAAGAGCTG GTCTCTGCTG CCAGTAAGTA
420

TTTATGTTAC TAATTAACTG TGTTGTCCTA GTTTCTTAAT GTTTACTGCA ATAAGCCTAG
480

AAAATTGTTT GAGGGGAAGT GATTGAGGGC ACAGAAACCT AAAACACATA CACAAATTAT
540

GCACAACTGC CAAATGAAAG TATTCTTGCT TGCTGTCTAA CTCAANAATT CTATTATTTT
600

T
601

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 387 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 18
(B) LOCATION: 80..185

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 25..44

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 218..237

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GTGCATATAT ATGTGAGAAT TTTGCTCAAT CCAGTAGCCC AGAAAGCCAA ACCATTTATC
60

TCTTACTGTT CTATCCCAGA GTCTGACAAC CTGGATTACC GCCTAGGAGC TATTCACTCC
120

CTGGTATATA AGCTACCAGA AAAGAACCGA GAGATGCTGG AACTTCTGAT AAGACACTTG
180

GTCAAGTAAG TAACTGCTGG ATTTTCAGAA AAAGTTCCTA TTAGAGGACT GGCCCATGTG
240

GTTGGACTAC ACAGAAACTG CCTCTCAGCT CTTTCAGCCC CAGCCCTTAA GTGCTTCCTT
300

GGAAGCTGAA TGCTCTGTGA GGAAGGCTAT TTTGCCTTGA CCCATGTACA TATCCTCTTA
360

GAGTCATCAT GCATGTGGAT TGTCTCA
387

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 460 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 19
(B) LOCATION: 79..238

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 51..70

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 252..271

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AACAGCACCT AAAACAGTCT TGGTTGTAAG GGGATACTGG AGCAAATTTT GTTAATCTTG
60

CCCCTTTTCT TCTGGCAGTG TGTGTGAGCA CAGCAAAGAG AATCTTATGA CCCCCTCCAA
120

CATGGGAGTA ATCTTTGGGC CCACCCTGAT GAGAGCTCAG GAGGACACTG TGGCCGCCAT
180

GATGAACATC AAATTCCAGA ACATAGTGGT GGAAATACTA ATCGAGCACT TTGGCAAGGT
240

ATGCATTTTC TATTCTCACT ACCTGTCTTC CAAACATGTG ACACTTTCCC CCAACTGCCT
300

TTTAGTGCTG TGTCTTCCTC CTTGGCTCAC GTTGACAGTG AAAGGAAATC CCATTATGAC
360

ACAATGACAT TTAATGGCAA CTCTGACCCT GGGAAATTCA TTCATTCAGC AAACATTGCT
420

TAAGCTTATA ACTATATTAT TTTCAGACAC CATGCTAAAT
460

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 512 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 20
(B) LOCATION: 230..377

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 206..225

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 383..402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TCTTCCAAAC TGAGAAGTGC CAGGTTNTGT GCCTTGAGCA TAGTAGGAGN TACNTAAACA
60

TTTACCTGTA GNTAGAGTGA TTAAGAAAAT CTCTGATTCT TTGAGTCATG TTAGTATTCA
120

CGTNACAAAC TCTAGATATA AGGCCAACAA GCATCAANTG GTGGGTAGCA TTCAGAAGAC
180

AAAAANTTGA TNTAANTATT CTNTAGATAT NTTCCTTCTT TNTCCACAGA TCTATTTAGG
240

TCCACCTGAG GAAAGCGCTG CACCGCCAGT GCCTCCGCCT CGGGTGACAG CAAGAAGGCA
300

CAAACCAATC ACGATTTCAA AGCGCTTGCT GCGAGAAAGG ACGGTTTTCT ATACTTCTTC
360

CCTGGATGAA AGCGAAGGTC AGTACTNAGG TTCTCCTTTA GCTTCTGAAT GGTGATTAGA
420

CACNNAGNAN GATATCNAAT GGCTCAAGCG GTGGCATCAC CATTTNTCTC TCTATAAAAG
480

TANACCTTTC CTGNCTCCTG AACTTAAAAG CA
512

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 841 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 21
(B) LOCATION: 185..508

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 151..170

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 511..530

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGCTTTACAT GAACTGGGAA GGGTAGAGAA TGATTTTGTG GGATATAGTT GGTTTGTGCC
60

ACAGTGACAT AACTGCTTTG AAAATGTATA CAAATTTTCA AAATTAAGTA TGTATGCATG
120

TATCAAAATG AAAAGGTTTT AAAAGTTATC ATTAATCTTC CCTNTTGGCA CCAACTTTTC
180

CTAGATGAAA TCCAACATCA AACACCGAAT GGTACTATCA CCAGCAGCAT AGAACCCCCC
240

AAGCCACCAC AACACCCCAA ACTACCTATT CAGAGGAGTG GGGAAACTGA TCCTGGGAGG
300

AAGTCCCCAA GCAGGCCTAT TTTGGATGGC AAGTTGGAGC CCTGCCCAGA GGTGGACGTG
360

GGGAAGTTGG TGTCTAGGCT GCAGGATGGA GGGACCAAGA TCACCCCAAA GGCCACCAAT
420

GGACCCATGC CAGGCTCTGG GCCCACCAAG ACCCCCTCTT TCCACATAAA GAGACCAGCT
480

CCCCGGCCCC TGGCCCACCA CAAGGAGGGT AAGTGCTTGG GAATCCCATG GGAGCCAGAG
540

CTGACCCTAA CTACTTTTCA CCTTGAGATC CTTCTGAGTT TGGAGATATA TTTAAGTGGA
600

AATATGTTCC AGTTTACTAC CACTAATATT GGAACAGTGG GCAAGATCAC AATAATCAGT
660

CACAATAATC ACTAGAATGT AAGCTCCATG AGGGCCGGGA TTTTTTACCT GTTTTGTTGA
720

CCTCTATATC CCAAGTGCTA TGTGCCTGGC ACTGTACTAA TTGCTGATAT ACTATTTCTT
780

ATCCTCACAA TCCCACTGTA AAGAATGTAT TATTCTTAAT ATTTTCTTTT TTTTTTTTTT
840

T
841

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 615 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 22
(B) LOCATION: 320..485

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 294..313

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 496..515

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TTTAATCCTC CCACTATCTC TGTAAGATAA TATATTGTGG ATCTTTATTA TATAGCTGGG
60

GAAACTGAGA CTTAGGGAAT GGATATGACA CACCCAAGAT ATNTGAAACT CCAGAGCTGG
120

GGTTCAAATA TAGACTTTCT GAAGGGACAG TTGCCAGAAA AATTACAAAA AAAAAAAAAA
180

ATAGCCAGAG TTGTTAGTCA CCAAGAAGAA ATGGAGGCCA AGGAAGTTGG CCCAGGTAAC
240

TCTCATATTG GGTGCCTGCT CATGAGTAGT GTTCTGTTTG GCTAACCATC CAAGTTCCTG
300

GTATCATTTT CTCTTCCAGG GGATGCTGAC AGTTTCAGCA AAGTGCGGCC TCCAGGAGAA
360

AAGCCAACCA TCATCCGCCC CCCAGTGAGG CCCCCAGATC CTCCCTGCCG GGCAGCTACT
420

CCCCAAAAGC CAGAACCAAA GCCAGATATT GTGGCTGGCA ATGCGGGGGA AATCACATCA
480

TCTGTGTCAG TAGGGTTGTA CCTCAAAGTT GACTGAAGTC CTGTACTAGG CCACTAGGAA
540

TGCTTTCAGG ATCACCATAT TAAGGGTATA CAGTGCACAG CCCTGGGGCA TCCTTCACTT
600

TATAGTCTAG GGAAA
615

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 475 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 23
(B) LOCATION: 211..261

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 179..198

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 271..291

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

AATGGGGATA AAGAGTGTCA GCTATGGCCT TAGGGGTGCC TATGGGCTCT GGGCCATTTC
60

ACATTTGTAT GTGTAGGGCC TTTGCCAGCA AAGGCAGGGG CTGGCATTGG TGTCCCATCT
120

GGTTCAGAGT CTCCTGTCCT TTCTGTTGGC CATTGGTTCT CACGTGTATA CCAAAGCAAC
180

TTATGGGACT TGGTTGGCTT CTGTTTGCAG GGTGGCTTCC AGGACCAGGT TTTTTGAAAC
240

AGCTTCCCGG AAAACAGGAA GGTAAGATAT GGAGGTGACA AAAGAAAAAC CAAATCGCCT
300

TTTAATAACT GCATCCTTAG CATACAATTG TGCTCACTCT AACATCTTTC TCTTTTTGTT
360

TCTCTACAGC TCTGTCTCTG TCTGTCACTT TCTCTTCCCC AATTCTGTCT CTCCATCCCT
420

ATCTGTCTGT CACCTGTTCA CCTGTGTGTC TATTTGTTTC TCTCATATTC TTTTT
475

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 238 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 24
(B) LOCATION: 115..156

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 65..84

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 165..184

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CCATTTTTCA AATGGCCTCT TTAGCACTGG CCTAGAAGTG TCCCCCATTC CCCCAATTTA
60

CCTTTCCAGT CCTGATTTCT AGAATCTTAG TGAAACGTCT TTCTTTATCC ACAGTTCTCA
120

AGGCAGACTT CCTGGAGATG AAAGTTGAGG CTACAGGTAT GCAGTCCCCA TCCCTGATTA
180

CAAAATCTTG TTCCACATAA GCCTTCATTA CGGGATCTGA TATTTTGAGG ACTGGAAT
238

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4504 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: exon 25
(B) LOCATION: 1..4235

(ix) FEATURE:
(A) NAME/KEY: polyA_site
(B) LOCATION: 4236..4241

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GTTTTAAAAG CCTTGGCCTC AGAGGACCCT TTCCAGGTTC TGAAAGAGTC TTTCTTCTTC
60

AAACCTTTGT GTGCGGAGTC ATTTTGTGTT GAAGAGCAGC TCCTTCCTAG CCTTGCACTT
120

TCAGACTCTC TCTGGGAGGC CATAAAATAA GGAGCATATG TCCTAGACAG GTGTTTATAT
180

CTCCTTTGTA TTCTGTCTTC ATCCCCTCAG AAGGTCTGTT TTGAGTTCCT ATAACACTGT
240

GAAGAGCTGG ACTCCCTCAA GCCAGACTCT GCCAAAACCA AGATATCCAC TTACCTGAGT
300

TGAAGAGGGG AGCTCAGTTT TCAACTCTTC CCTGAACTTC CTGCTTCCTC AGAGGGCCAT
360

TGAACTCTGA GAGATTTGGG GCTAAAGACT GATCTCAGGG GTCTTACCTT GAACTGAAGG
420

CCACTTGAGT TGGGGCCATT GCTTACCTTG GTTGGAAGGG AATAGAAATG TTTGCTGAAC
480

ATTGGAGAAT CTCAACATGT CTCCTACTGA GGATATGGAC ACTGGTGCCA TGTCAGCGCT
540

CTGGTGCTGC AGTATGTTGC CAAGAGCCCG TCTGCTCCTG CGAGGCTATG AGTGGGATGA
600

GTGATGCCCC CACAGCACCT CCATGTGGAC TTAGGAAGGT GGCCTTCCTG CTGTTACATG
660

CAGCCACTTA GGACAAATCT GCAAAGCATG TTTTGCATGT AAAAGCCTAG GTCTATTTGG
720

ATTATTCTTT CTCCTTTTTT TTGACAGCTT CCTGTCAAGC AATCAAGAAA CAAACAAAAG
780

CTGAACACAT TTCCTTTTAA AAAAAGGAGA CTGTGTTTGG TCCTGTAGGA GTTCTATTTT
840

GGGGTCAAAT GCTAGAAAAA TTGTTAAGGT GGATTGAGGC CAGGCAGCTG TCACTGCTGC
900

TTCATGTTTG CCTTCTGCAC ATAAACTCTT TTATCTCCTG AAAAAAGCAG TTCTTAACCC
960

AGTGTCCATG GACTCAGAAA CTCCATGATG CCCCTGAGAT GGTATGCACA ATTCCATGAC
1020

AATATGCCCT TTCTGGGGAG ATAGTCCATA ATGTTCTGCT AAATTTCAAA TGGGCTCGTG
1080

ACCCAAAAAA GTCAAGAACC ACAGCACTTG AGTTAAAATA CTCTTTTTAC AATCCATATA
1140

AGCCCTTGAT TGGAAGGGCT TTTCAAAATC ATTTAGTCTA ACAACTGCCC AGTTTCCAGT
1200

CGGGGGAACT GAGGCAGAGC AAGGTAGTGA TCACACCAGT ACAAGATTTC AGGTCCCAGG
1260

CTCCTATGCA AGTTTTTTTT CCCCATTATA TCACACTTAT TTAGCAAGGG ACCTTGTGGT
1320

TTGTGGCTTT AGTGGCCATC ATTTCTGGGG GTTGGCTTTT ACCCTTTTTC TTGAATATTT
1380

GCCACCAAGT GAAAAATGTT AGGACATAAA CCCTTGCCAG GTCCCTTTCA TTTGCTATCT
1440

CTATTATTGG AAAGGACCTA AAAATTGGTG TAATGGGGCA GAAATCTGAG GAATGGACAT
1500

TTCTAATTCC TGTTTGTTGA AGGGAAGTTG CTGGAAAGAG CATCAGTACT TGTTTCTATG
1560

CAGATGCCTG GGCCGTAGCT TGTCTGTAGC GTCTGTATAA TTATAATGTT GCCCAGTGTG
1620

AGGGAAAGAG CTTTCCTACT TGCACTCTTC TACCAAGGCC CTGTTAGTGC ACTGATTATA
1680

GTACTGACAG ATAAAGCCTA GATGAGAGAG ATAGAGAGTG AGTACATGCA CACTCATGTG
1740

CAAACCCACT CAGAGATGCA TTTGGAACAG TGCTACTGAA AGGCAGTAGT CATTTTCAAG
1800

ACTGAATTCC AAACATGGTT TATTGGTGAG TTAGGAACAT GTAAGGCCAA GTACACTGAG
1860

AGCCTTTTTG AAAGTAATTG AGTGGAAACT TGATGCCATT CTAAATCAAG GCATATCCAG
1920

GTGGCCCGGT TTGAACTCAC TCCACTGTAC CCAGTCTCAA AGGCCAGGTT GCTAAGAAAC
1980

CAGGAGTAAA AGAGTCAAGT GACCATCATT TCACCTGCTG CTTGCCCCCA ATAGTAGTCT
2040

CTGTGAGGCC TTACTGACCT CACCTAGGAA GTGATTTTTG AGCCCTTGTT TCAGGGCTGT
2100

GGCCTCCCTG CTCTATCCTG AATAAAGCAG ACAGGTGTGC AGATTTTGGC CATGAAAGCA
2160

TGGCTAATAG GGCCACAGTC CCTTTAAAGA AACATGGTTT GACTCTGGTT TTCTTGGGGG
2220

AAAATACCAC AATCACCGAT GCAAACATTG GAAGATTATT GAGAGCCCTA GAAAGCTGCT
2280

GTGATCCCAG TAGAAAATAT GTCCCAGAAA TGTCATGAGA TTGCTGTGTG TTGCCTGGGA
2340

CACAGATCAA GGGCCTATCT TGGAGAGCTG GGGACCAGCA GTCTGCCTGG AGGCCAGGGA
2400

GCAGTGGCTG AGTAGCTCTG CCTTTGCTCT GGTCTATACC TTAAGAATGC CAAAGAATGA
2460

ATTTCAACGC CTGCCTTTGG CACTCTGACT TAAAGTGCAA AAAGCTTCTG TGGCGAGGCA
2520

TGCTATCATG GAATGAGACT GGCTTGCCCT AGGCTTAATG GATGGGCAGT CATTTTGCAG
2580

AGGCTATGGG AAGAGGGTGA TAATAGAAGA GTGGCAGCTA TAGGAAATTA TCAACATACC
2640

TTGGCCAGCA AGTTAGAGAA TCTGGCAATG GATGAACTGA AAGTGATGAA CTGGCAGGGA
2700

TAACAAAGAA CCTAACATTT ATTAAGCACG TATTTATTAA CTGCTCAGTG TTTCATATTC
2760

ATGCAAGTAT TCTCATTTTA CAGAGAAAGA AATTATGGCC CAGGGGGCTA AAGTAAACAA
2820

CTCAAGGGCA CATAGAAAGT AAATAAAAGG ACTGTGATTT GAATCCAGGC CACTCTTAGC
2880

CCATGCTGTT TTCCCTTTGC CACACTGTGG TAGGTGTTTG AACAGAGGCC ACATTACTAG
2940

AGTTGGCATG ACTCTTGACT CTTGCCTGCC TAACAAAATA TTGAAAGGCA AACATTTGAA
3000

GGAGGGAGGG GGTGCAGGTT CAGTTTATAT GGAAATGCAA ACTGGGCTGG AAGATATTCC
3060

TGAGTTAGGC TTTCTCTTCA TATTCAGCTT GCACATTTGG TAATGTTTTT AAAATGATCA
3120

TCTAATTTTA TTTTGTGAAG TGAAGGATTT GTGTTTTAGT TGGCAGTTGT TAAGTCCTTG
3180

GCTTGCCATT TTTCAAAAAG TAAAAAGGTC CTCACAGGTG TCTCCATACT TCGCCAAGGT
3240

TGTAGCATGG GCAGTTTCAG TTTCAGCCTA AGAGACTGGT GACATCCACA AATGCAGTTT
3300

TAGAAGCAGA AAAGGTCTTG GTGCCTCTGC AGTACTTGAT GTATTGGGGT CAAATCTCTA
3360

CAAATTTTTC TGTGGTGATA GCAAAATCAA GAGATGGCTT ACAAAAAGAA ATATTGAATT
3420

TTTATTTTTG AAGTTTTTGT TTTTTAAAAG GTTGGGGGTG TTCAGCCACT GAGGGACAAA
3480

ACTTAGCATC TAATTTCAAT TATAGTGTCA TGCAGAGTAT TTCTAAAGTA ATTGGTTATC
3540

ATGGGAAAGT ATTCTCTTTT CAAGAAGTTC TTTGATTCTG TAATAACTAG AACAAATAAA
3600

GTAGTAAAAG AAGAAATAGT TCTGTGACTA GGAAAAAATT GCTTTTGAGA GAACATAGAT
3660

CAATTATACT ACTTCTAAGG TAGCTGCAGA TAAGTGGCCT TGACACATTA CAAGCCTGGA
3720

AAAAAACATC AGAAATAATA AAAAATTTCA GAGAGAATCA AGATACCTTT TTTTTTCTTT
3780

TTTTTTTCTT TTTTTTATTA TACTCTAAGT TTTAGGGTAC ATGTGCACAT TGTGCAGGTT
3840

AGTTACATAT GTATACATGT GCCATGCTGG TGCGCTGCAC CCACTAATGT GTCATCTAGC
3900

ATTAGGTATA TCTCCCAGTG CTATCCCTCC CCCCTCCCCC GACCCCACCA CAGTCCCCAG
3960

AGTGTGATAT TCCCCTTCCT GTGTCCATGT GATCTCATTG TTCAATTCCC ACCTATGAGT
4020

GAGAATATGC GGTGTTTGGT TTTTTGTTCT TGCGATAGTT TACTGAGAAT GATGGTTTCC
4080

AATTTCATCC ATGTCCCTAC AAAGGATATG AACTCATCAT TTTTTATGGC TGCATAGTAT
4140

TCCATGGTGT ATATGTGCCA CATTTTCTTA ATCCAGTCTA TCATTGTTGG ACATTTGGGT
4200

TGGTTCCAAG TCTTTGCTAT TGTGAATAGT GCCGCAATAA ACATACGTGT GCATGTGTCT
4260

TTATAGCAGC ATGATTTATA CTCATTTGGG TATATACCCA GTAATGGGAT GGCTGGGTCA
4320

AATGGTATTT CTAGTTCTAG ATCCCTGAGG AATCGCCACA CTGACTTCCA CAATGGTTGA
4380

ACTAGTTTAC AGTCCAACCA ACAGTGTAAA AGTGTTCCTA TTTCTCCGCA TCCTCTCCAG
4440

CACCTGTTGT TTCCTGACTT TTTAATGATT GCCATTCTAA CTGGTGTTGA GATGATATCT
4500

CATA
4504

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 639..3047

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 152..172

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 586..606

(ix) FEATURE:
(A) NAME/KEY: oligonucleotide
(B) LOCATION: 641..663

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TGTGCGTCGC GCTCTCGCCC TCCTCTTCCC GCTGCAGTGT CTATGGAGCG AGGCTACGTT
60

TCATTGCCGC CCTGGCTTAA CCCTTCCGGC GCCTAAAAGG ACGGCCGGCC GGCCGGTCCC
120

TTGCACCAGG AAGAAGTCTT AGCAGCCAGC GGGCCCTGGT CAGGAAACTC TAAGGTACAA
180

GGAAAACAGT TGAGGAAGGA GCCAGAGCGC TCCGGTTTGG TCCTCGGGCT TCGCTGGGGC
240

GGGGCGCAGG CGTTGGCTTT AAGAAAGGGG AGGGGACAGT GCAATCCGGG TTGCCCGCGG
300

ATTCGGCCAA GGAATCTTCC GCTCGCTCCG GAGCGAGGAG CCTGTAAAGA GGCTGTTCCC
360

AGCTCCAGCT CTAACCTCGC CTACACCTTG GGCGGGCCCA ATGTCACGTT TGCAATTGCT
420

CAGGAAGGAT CCGGCCCGTC TCCGGAGGCA AGTCGGGCTG CGGTTTTTGC TGCTTATCTG
480

GGAAGGCGAT GCCTAAGGGA CATGCTGCTT GCTAGGCAGC ACCCTGCCGG GATCCGACTG
540

CGATAGTTAG CTCTCCCTGG CCCTGAAGCC ATCGCCGGGG CGCCTGTTCT CTGTCCGGAC
600

CAGCCAGCGC TCCTCAGGAG TCTCACTGAA ACAGAACCAT GGGTCATCCC CCGCTGGAGT
660

TCAGCGACTG CTACCTGGAC AGCCCCGATT TCCGCGAGAG GCTCAAGTGT TATGAGCAGG
720

AACTGGAGAG GACCAACAAA TTCATCAAAG ACGTAATCAA AGACGGCAAC GCGCTTATCA
780

GCGCTATGAG AAATTATTCT TCTGCTGTTC AGAAATTTTC CCAGACGCTG CAGTCATTTC
840

AGTTTGATTT CATTGGAGAC ACTCTGACTG ATGATGAAAT TAACATCGCT GAATCCTTCA
900

AGGAATTTGC TGAATTGCTC AACGAGGTAG AAAATGAGAG GATGATGATG GTACACAATG
960

CTAGTGATTT GCTGATTAAA CCCTTGGAAA ATTTCCGGAA GGAACAAATA GGCTTCACCA
1020

AGGAGCGGAA AAAGAAATTT GAAAAGGATG GTGAGAGGTT TTATTCTTTA CTGGATCGGC
1080

ACTTACACCT GTCTTCAAAA AAGAAAGAAT CTCAGTTACA AGAGGCAGAC CTACAGGTGG
1140

ACAAGGAGAG GCACAATTTT TTCGAGTCCT CTCTTGATTA TGTTTATCAA ATCCAGGAAG
1200

TTCAGGAGTC CAAGAAGTTC AATATTGTGG AGCCTGTCTT GGCCTTTCTT CATAGTCTGT
1260

TCATTTCTAA CAGCCTGACT GTGGAGCTCA CACAGGATTT CCTCCCATAC AAACAACAGC
1320

TCCAACTCAG TTTACAGAAT ACAAGAAATC ATTTCTCCAG TACCCGGGAA GAGATGGAAG
1380

AACTTAAGAA AAGGATGAAA GAAGCTCCCC AGACATGCAA ACTTCCAGGA CAGCCAACTA
1440

TTGAAGGCTA TCTCTATACA CAAGAGAAAT GGGCTTTAGG AATATCCTGG GTGAAATACT
1500

ATTGCCAGTA TGAGAAAGAG ACCAAAACAC TGACCATGAC GCCTATGGAG CAGAAGCCAG
1560

GTGCTAAGCA GGGGCCCTTG GACTTAACAC TGAAGTACTG TGTGAGAAGG AAGACGGAGT
1620

CTATCGACAA GAGGTTCTGT TTTGACATAG AAACTAATGA AAGGCCAGGA ACCATCACTC
1680

TGCAGGCCCT TTCAGAAGCT AACAGAAGGC TATGGATGGA AGCCATGGAT GGGAAAGAAC
1740

CTATCTACCA CAGCCCTATA ACAAAACAGC AAGAAATGGA GCTAAATGAA GTGGGCTTCA
1800

AGTTTGTCAG GAAGTGCATC AATATTATTG AGACCAAAGG GATCAAGACA GAAGGGTTGT
1860

ACCGCACTGT GGGCAGCAAT ATTCAGGTTC AGAAGCTGCT GAATGCCTTT TTTGATCCTA
1920

AATGCCCAGG AGATGTTGAT TTTCATAATA GTGACTGGGA CATTAAGACA ATCACCAGCT
1980

CCTTGAAATT CTACCTCAGG AATCTTTCTG AACCTGTCAT GACCTATAGA CTTCACAAAG
2040

AGCTGGTCTC TGCTGCCAAG TCTGACAACC TGGATTACCG CCTAGGAGCT ATTCACTCCC
2100

TGGTATATAA GCTACCAGAA AAGAACCGAG AGATGCTGGA ACTTCTGATA AGACACTTGG
2160

TCAATGTGTG TGAGCACAGC AAAGAGAATC TTATGACCCC CTCCAACATG GGAGTAATCT
2220

TTGGGCCCAC CCTGATGAGA GCTCAGGAGG ACACTGTGGC CGCCATGATG AACATCAAAT
2280

TCCAGAACAT AGTGGTGGAA ATACTAATCG AGCACTTTGG CAAGATCTAT TTAGGTCCAC
2340

CTGAGGAAAG CGCTGCACCG CCAGTGCCTC CGCCTCGGGT GACAGCAAGA AGGCACAAAC
2400

CAATCACGAT TTCAAAGCGC TTGCTGCGAG AAAGGACGGT TTTCTATACT TCTTCCCTGG
2460

ATGAAAGCGA AGATGAAATC CAACATCAAA CACCGAATGG TACTATCACC AGCAGCATAG
2520

AACCCCCCAA GCCACCACAA CACCCCAAAC TACCTATTCA GAGGAGTGGG GAAACTGATC
2580

CTGGGAGGAA GTCCCCAAGC AGGCCTATTT TGGATGGCAA GTTGGAGCCC TGCCCAGAGG
2640

TGGACGTGGG GAAGTTGGTG TCTAGGCTGC AGGATGGAGG GACCAAGATC ACCCCAAAGG
2700

CCACCAATGG ACCCATGCCA GGCTCTGGGC CCACCAAGAC CCCCTCTTTC CACATAAAGA
2760

GACCAGCTCC CCGGCCCCTG GCCCACCACA AGGAGGGGGA TGCTGACAGT TTCAGCAAAG
2820

TGCGGCCTCC AGGAGAAAAG CCAACCATCA TCCGCCCCCC AGTGAGGCCC CCAGATCCTC
2880

CCTGCCGGGC AGCTACTCCC CAAAAGCCAG AACCAAAGCC AGATATTGTG GCTGGCAATG
2940

CGGGGGAAAT CACATCATCT GTGGTGGCTT CCAGGACCAG GTTTTTTGAA ACAGCTTCCC
3000

GGAAAACAGG AAGTTCTCAA GGCAGACTTC CTGGAGATGA AAGTTGAGGC TACAGGTTTT
3060

AAAAGCCTTG GCCTCAGAGG ACCCTTTCCA GGTTCTGAAA G
3101

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 802 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Met Gly His Pro Pro Leu Glu Phe Ser Asp Cys Tyr Leu Asp Ser Pro
1 5 10 15

Asp Phe Arg Glu Arg Leu Lys Cys Tyr Glu Gln Glu Leu Glu Arg Thr
20 25 30

Asn Lys Phe Ile Lys Asp Val Ile Lys Asp Gly Asn Ala Leu Ile Ser
35 40 45

Ala Met Arg Asn Tyr Ser Ser Ala Val Gln Lys Phe Ser Gln Thr Leu
50 55 60

Gln Ser Phe Gln Phe Asp Phe Ile Gly Asp Thr Leu Thr Asp Asp Glu
65 70 75 80

Ile Asn Ile Ala Glu Ser Phe Lys Glu Phe Ala Glu Leu Leu Asn Glu
85 90 95

Val Glu Asn Glu Arg Met Met Met Val His Asn Ala Ser Asp Leu Leu
100 105 110

Ile Lys Pro Leu Glu Asn Phe Arg Lys Glu Gln Ile Gly Phe Thr Lys
115 120 125

Glu Arg Lys Lys Lys Phe Glu Lys Asp Gly Glu Arg Phe Tyr Ser Leu
130 135 140

Leu Asp Arg His Leu His Leu Ser Ser Lys Lys Lys Glu Ser Gln Leu
145 150 155 160

Gln Glu Ala Asp Leu Gln Val Asp Lys Glu Arg His Asn Phe Phe Glu
165 170 175

Ser Ser Leu Asp Tyr Val Tyr Gln Ile Gln Glu Val Gln Glu Ser Lys
180 185 190

Lys Phe Asn Ile Val Glu Pro Val Leu Ala Phe Leu His Ser Leu Phe
195 200 205

Ile Ser Asn Ser Leu Thr Val Glu Leu Thr Gln Asp Phe Leu Pro Tyr
210 215 220

Lys Gln Gln Leu Gln Leu Ser Leu Gln Asn Thr Arg Asn His Phe Ser
225 230 235 240

Ser Thr Arg Glu Glu Met Glu Glu Leu Lys Lys Arg Met Lys Glu Ala
245 250 255

Pro Gln Thr Cys Lys Leu Pro Gly Gln Pro Thr Ile Glu Gly Tyr Leu
260 265 270

Tyr Thr Gln Glu Lys Trp Ala Leu Gly Ile Ser Trp Val Lys Tyr Tyr
275 280 285

Cys Gln Tyr Glu Lys Glu Thr Lys Thr Leu Thr Met Thr Pro Met Glu
290 295 300

Gln Lys Pro Gly Ala Lys Gln Gly Pro Leu Asp Leu Thr Leu Lys Tyr
305 310 315 320

Cys Val Arg Arg Lys Thr Glu Ser Ile Asp Lys Arg Phe Cys Phe Asp
325 330 335

Ile Glu Thr Asn Glu Arg Pro Gly Thr Ile Thr Leu Gln Ala Leu Ser
340 345 350

Glu Ala Asn Arg Arg Leu Trp Met Glu Ala Met Asp Gly Lys Glu Pro
355 360 365

Ile Tyr His Ser Pro Ile Thr Lys Gln Gln Glu Met Glu Leu Asn Glu
370 375 380

Val Gly Phe Lys Phe Val Arg Lys Cys Ile Asn Ile Ile Glu Thr Lys
385 390 395 400

Gly Ile Lys Thr Glu Gly Leu Tyr Arg Thr Val Gly Ser Asn Ile Gln
405 410 415

Val Gln Lys Leu Leu Asn Ala Phe Phe Asp Pro Lys Cys Pro Gly Asp
420 425 430

Val Asp Phe His Asn Ser Asp Trp Asp Ile Lys Thr Ile Thr Ser Ser
435 440 445

Leu Lys Phe Tyr Leu Arg Asn Leu Ser Glu Pro Val Met Thr Tyr Arg
450 455 460

Leu His Lys Glu Leu Val Ser Ala Ala Lys Ser Asp Asn Leu Asp Tyr
465 470 475 480

Arg Leu Gly Ala Ile His Ser Leu Val Tyr Lys Leu Pro Glu Lys Asn
485 490 495

Arg Glu Met Leu Glu Leu Leu Ile Arg His Leu Val Asn Val Cys Glu
500 505 510

His Ser Lys Glu Asn Leu Met Thr Pro Ser Asn Met Gly Val Ile Phe
515 520 525

Gly Pro Thr Leu Met Arg Ala Gln Glu Asp Thr Val Ala Ala Met Met
530 535 540

Asn Ile Lys Phe Gln Asn Ile Val Val Glu Ile Leu Ile Glu His Phe
545 550 555 560

Gly Lys Ile Tyr Leu Gly Pro Pro Glu Glu Ser Ala Ala Pro Pro Val
565 570 575

Pro Pro Pro Arg Val Thr Ala Arg Arg His Lys Pro Ile Thr Ile Ser
580 585 590

Lys Arg Leu Leu Arg Glu Arg Thr Val Phe Tyr Thr Ser Ser Leu Asp
595 600 605

Glu Ser Glu Asp Glu Ile Gln His Gln Thr Pro Asn Gly Thr Ile Thr
610 615 620

Ser Ser Ile Glu Pro Pro Lys Pro Pro Gln His Pro Lys Leu Pro Ile
625 630 635 640

Gln Arg Ser Gly Glu Thr Asp Pro Gly Arg Lys Ser Pro Ser Arg Pro
645 650 655

Ile Leu Asp Gly Lys Leu Glu Pro Cys Pro Glu Val Asp Val Gly Lys
660 665 670

Leu Val Ser Arg Leu Gln Asp Gly Gly Thr Lys Ile Thr Pro Lys Ala
675 680 685

Thr Asn Gly Pro Met Pro Gly Ser Gly Pro Thr Lys Thr Pro Ser Phe
690 695 700

His Ile Lys Arg Pro Ala Pro Arg Pro Leu Ala His His Lys Glu Gly
705 710 715 720

Asp Ala Asp Ser Phe Ser Lys Val Arg Pro Pro Gly Glu Lys Pro Thr
725 730 735

Ile Ile Arg Pro Pro Val Arg Pro Pro Asp Pro Pro Cys Arg Ala Ala
740 745 750

Thr Pro Gln Lys Pro Glu Pro Lys Pro Asp Ile Val Ala Gly Asn Ala
755 760 765

Gly Glu Ile Thr Ser Ser Val Val Ala Ser Arg Thr Arg Phe Phe Glu
770 775 780

Thr Ala Ser Arg Lys Thr Gly Ser Ser Gln Gly Arg Leu Pro Gly Asp
785 790 795 800

Glu Ser

" A new gene called oligophrenin 1, its expression product, and the diagnostic and therapeutic applications thereof".

The present invention relates to the identification of a new gene, called oligophrenin 1, and its expression product, as well as to the diagnostic and therapeutic applications of these nucleotide and peptide sequences.

A major challenge for human genetics is the identification of new causes of mental retardation, which, although present in about 3 % of individuals, is unexplained in over half of all cases. X-linked mental retardation is acknowledged to be a major cause of severe learning difficulties. Surveys have shown an excess of males over females with severe mental retardation and later studies suggested that the excess was the result of an X-linked condition. X-linked mental retardation (XLMR) is a vastly heterogeneous group of disorders which can be roughly categorized as syndromic (MRXS) or non specific (MRX). Families with syndromic disorders usually have a quite distinct phenotypic presentation whereas families with non specific disorders present no distinctive somatic features. Despite recent advances in identifying genes such as FMR1(Verkerk et al., 1991), FRAXE (Knight et al., 1993 ; Gecz et al., 1996), L1-CAM (Vits et al., 1994), FGD1 and XH2 (Gibbons et al., 1995), involved in MRXS conditions, so far no gene significantly involved in MRX has yet been identified or cloned. Compilation of the literature and McKusick's catalogue data revealed at least 95 X-linked disorders in which mental retardation appears as the main feature. Of these 95, 40 have been regionally mapped on the X chromosome by conducting linkage studies using DNA markers in single large families or in a collection of families with the same XLMR syndrome. Several loci appear to be located in the proximal Xq region. However, it is impossible to evaluate how many MR genes there are in reality, partly because of the broad localisation and the presence of several overlaps between intervals of assignment. Thus, fine mapping and identification of genes implicated in nonspecific X-linked MR essentially depend on thorough investigation of molecular abnormalities such as balanced translocation, inversion or contiguous gene deletion associated with MR.

Bienvenu *et al.* (1997) have recently reported a molecular cytogenetic investigation of an X;12 balanced translocation observed in a female affected with a mild mental retardation and have localised the breakpoint in Xq12.

The authors of the present invention have now cloned the gene responsible for MRX, which they have called the oligophrenin 1 gene.

In order to define the genomic structure of this gene, the authors of the present invention constructed and investigated a cosmid/phage contig that covers the gene. Determination of exon-intron boundaries was performed through sequence comparison between cDNA clones and genomic DNA, which led to the identification of 25 exons.

The authors of the present invention have thus isolated and characterized the oligophrenin 1 transcripts. Said transcripts contain an open-reading frame (ORF) which is encoded by exon 2 to exon 24. This ORF is 2406 bases long and encodes a protein of 802 amino acids, called the oligophrenin 1 protein.

A subject of the present invention is thus an isolated nucleic acid having a sequence selected from the group consisting of sequence SEQ ID n° 1 to SEQ ID n° 25, and a homologous nucleotide sequence thereof.

SEQ ID n° 1 represents the 5' fragment of the genomic DNA of the human oligophrenin 1 gene.

SEQ ID n° 2 to SEQ ID n°25 represent fragments of the genomic DNA of the human oligophrenin 1 gene including exons as shown in table 1.

Table 1 : identification of the sequences

Sequence Identification	exon included in the fragment of the genomic DNA of the oligophrenin 1 gene
SEQ ID n° 2	exon 1 and exon 2
SEQ ID n° 3	exon 3
SEQ ID n°4	exon 4
SEQ ID n°5	exon 5
SEQ ID n°6	exon 6
SEQ ID n°7	exon 7
SEQ ID n°8	exon 8
SEQ ID n°9	exon 9
SEQ ID n°10	exon 10
SEQ ID n°11	exon 11
SEQ ID n°12	exon 12
SEQ ID n°13	exon 13
SEQ ID n°14	exon 14
SEQ ID n°15	exon 15
SEQ ID n°16	exon 16
SEQ ID n°17	exon 17
SEQ ID n°18	exon 18
SEQ ID n°19	exon 19
SEQ ID n°20	exon 20
SEQ ID n°21	exon 21
SEQ ID n°22	exon 22
SEQ ID n°23	exon 23
SEQ ID n°24	exon 24
SEQ ID n°25	exon 25

SEQ ID n° 26 represents the cDNA fragment corresponding to the common open-reading frame (ORF).

A subject of the present invention is also an isolated nucleic acid having a sequence selected from the group consisting of exon 1 to exon 25 as identified in the sequence listing and in table 2, and a homologous nucleotide sequence thereof.

Table 2 : identification of exon sequences

exon	from nucleotide n°	to nucleotide n°	Sequence which includes said exon
exon 1	1	634	SEQ ID n° 2
exon 2	778	935	
exon 3	403	498	SEQ ID n° 3
exon 4	483	544	SEQ ID n°4
exon 5	451	522	SEQ ID n°5
exon 6	416	517	SEQ ID n°6
exon 7	464	574	SEQ ID n°7
exon 8	244	348	SEQ ID n°8
exon 9	134	263	SEQ ID n°9
exon 10	383	483	SEQ ID n°10
exon 11	107	198	SEQ ID n°11
exon 12	211	289	SEQ ID n°12
exon 13	212	245	SEQ ID n°13
exon 14	172	234	SEQ ID n°14
exon 15	207	281	SEQ ID n°15
exon 16	270	354	SEQ ID n°16
exon 17	355	413	SEQ ID n°17
exon 18	80	185	SEQ ID n°18
exon 19	79	238	SEQ ID n°19
exon 20	230	377	SEQ ID n°20
exon 21	185	508	SEQ ID n°21
exon 22	320	485	SEQ ID n°22
exon 23	211	261	SEQ ID n°23
exon 24	115	156	SEQ ID n°24

"A homologous nucleotide sequence" is understood as meaning a sequence which differs from the sequences to which it refers by mutation, insertion, deletion or substitution of one or more bases.

Preferably, such homologous sequences show at least 70 % of homology, preferably 80 % of homology, more preferably 90 % of homology with any of sequences SEQ ID n° 1 to SEQ ID n° 26.

A polynucleotide of the invention, having a homologous sequence, hybridizes to the sequences to which it refers (any of sequences SEQ ID n° 1 to SEQ ID n° 26), preferably under stringent conditions. Parameters that define the conditions of stringency depend upon the temperature at which 50 % of annealed strands separate (T_m).

For sequences comprising more than 30 nucleotides, T_m is calculated as follows :

T_m = 81.5 + 0.41 (% G + C) + 16.6 Log (positive ion concentration) – 0.63 (% formamide) – (600/polynucleotide size in base pairs) (Sambrook et al, 1989).

For sequences comprising less than 30 nucleotides, T_m is calculated as follows :

$T_m = 4(G + C) + 2 (A + T)$.

Under appropriate stringent conditions avoiding the hybridization of non specific sequences, hybridization temperature is around from 5°C to 30°C, preferably from 5°C to 10° C below the calculated T_m, and hybridization buffer solutions that are used are preferably solutions with high ionic strength, such as an aqueous 6 X SSC solution for example.

A nucleotide sequence homologous to the open-reading from SEQ ID n° 26 means a nucleotide sequence which differs from the sequences to which it refers by mutation, insertion, deletion or substitution of one or more bases, or by the degeneracy of the genetic code so long as it codes for a polypeptide having the biological activity of oligophrenin 1 protein, as defined below.

Said homologous sequences include mammalian genes coding for the oligophrenin 1 protein, preferably of primate, cattle, sheep, swine, or rodent, as well as allelic variants.

The nucleic acid sequences of the invention are useful for the detection of an abnormality, such as a mutation, in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene. Such an analysis allows *in vitro* diagnosis of a neurological disorder associated with said abnormality.

A subject of the present invention is a method of *in vitro* diagnosis of a neurological disorder associated with an abnormality in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene, wherein one or more mutation(s), preferably inducing a modification of the expression of the oligophrenin 1 gene is detected in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene.

A subject of the present invention is also a nucleic acid comprising a sequence identical to SEQ ID n° 26 or to homologous sequences thereof, except for a one base deletion of the nucleotide 1578 as shown in SEQ ID n° 26.

The present invention relates to methods of *in vitro* diagnosis wherein the nucleic acid sequences of the invention or probes or primers derived thereof are used to detect aberrant synthesis or genetic abnormalities at the oligophrenin 1 gene level.

The present invention is more particularly directed to a method of *in vitro* diagnosis comprising the steps of :

- contacting a biological sample containing DNA with specific oligonucleotides permitting the amplification of all or part of the oligophrenin 1 gene, the DNA contained in the sample having being rendered accessible, where appropriate, to hybridization, and under conditions permitting a hybridization of the primers with the DNA contained in the biological sample ;
- amplifying said DNA ;
- detecting the amplification products ;
- comparing the amplified products as obtained to the amplified products obtained with a normal control biological sample, and thereby detecting a possible abnormality in the oligophrenin 1 gene.

The method of the invention can also be applied to the detection of an abnormality in the transcript of the oligophrenin 1 gene, by amplifying the mRNAs contained in a biological sample, for example by RT-PCR.

Thus another subject of the present invention is a method of in vitro diagnosis, as previously defined comprising the steps of :

- producing cDNA from mRNA contained in a biological sample ;
- contacting said cDNA with specific oligonucleotides permitting the amplification of all or part of the transcript of the oligophrenin 1 gene, under conditions permitting a hybridization of the primers with said cDNA ;
- amplifying said cDNA ;
- detecting the amplification products ;
- comparing the amplified products as obtained to the amplified products obtained with a normal control biological sample, and thereby detecting a possible abnormality in the transcript of the oligophrenin 1 gene.

This comparison of the amplified products obtained from the biological sample with the amplified products obtained with a normal biological sample can be carried out for example by specific probe hybridization, by sequencing or by restriction site analysis.

A subject of the present invention is also a nucleic acid sequence which specifically hybridizes with a nucleic acid sequence of the invention as previously defined or with their complementary sequences.

"A sequence which specifically hybridizes [...]" is understood as meaning a sequence which hybridizes with the sequences to which it refers under the conditions of high stringency (Sambrook et al, 1989). Such sequences are preferably oligonucleotides having at least 15, and more preferably at least 20 bases.

Such sequences, which are useful as primers or probes for the diagnosis methods according to the present invention may be preferably

selected from the group consisting of nucleic acid fragments of SEQ ID N° 2 to SEQ ID N° 26 as shown in table 3, or the complementary sequences thereof.

Table 3 : identification of the oligonucleotide sequences

Oligonucleotides		Sequence which includes said oligonucleotide
from nucleotide n°	to nucleotide n°	
727 958	746 977	SEQ ID n° 2
375 504	394 523	SEQ ID n° 3
418 551	437 570	SEQ ID n°4
423 553	445 574	SEQ ID n°5
388 540	407 559	SEQ ID n°6
436 584	458 603	SEQ ID n°7
219 363	239 381	SEQ ID n°8
108 336	128 355	SEQ ID n°9
361 492	380 511	SEQ ID n°10
81 223	100 242	SEQ ID n°11
188 300	207 319	SEQ ID n°12
166 259	189 278	SEQ ID n°13
133 250	152 269	SEQ ID n°14
151 293	170 315	SEQ ID n°15
221 363	244 382	SEQ ID n°16
305 438	324 457	SEQ ID n°17
25 218	44 237	SEQ ID n°18
51 252	70 271	SEQ ID n°19
206 383	225 402	SEQ ID n°20
151 511	170 530	SEQ ID n°21
294 496	313 515	SEQ ID n°22
179 271	198 291	SEQ ID n°23
65 165	84 184	SEQ ID n°24
152 586 641	172 606 663	SEQ ID n° 26

One skilled in the art very well knows the standard methods for analysing the DNA contained in a biological sample and for diagnosing a genetic disorder. Many strategies for genotypic analysis are available (Antonarakis et al., 1989, Cooper et al., 1991).

Preferably, one can use the DGGE method (Denaturing Gradient Gel Electrophoresis), or the SSCP method (Single Strand Conformation Polymorphism) for detecting an abnormality in the oligophrenin 1 gene. Such methods are preferably followed by direct sequencing. The RT-PCR method may be advantageously used for detecting abnormalities in the oligophrenin 1 transcript, as it allows to visualize the consequences of a splicing mutation such as exon skipping or aberrant splicing due to the activation of a cryptic site. This method is preferably followed by direct sequencing as well. The more recently developed technique using DNA chip can also be advantageously implemented for detecting an abnormality in the oligophrenin 1 gene (Bellis et al., 1997).

The cloning of the oligophrenin 1 gene, as well as the identification of various mutations responsible for neurological disorders according to the invention, allow direct or semi-direct diagnosis. The specificity and reliability of such diagnosis methods are more particularly appreciable for prenatal diagnosis. The nucleic acid sequences of the present invention represent a highly interesting tool for genetic counseling.

The authors of the present invention have shown that the oligophrenin 1 protein is a rho-GAP protein and that the constitutional loss of oligophrenin 1 activity in humans results in cognitive impairment. Defects in the oligophrenin 1 gene, or in the oligophrenin 1 gene product may cause inactivation of the oligophrenin 1 protein, which leads to constitutive activation of its target GTPases.

Such constitutive activation of rho family members has been shown to produce marked changes in certain actin-based processes, to alter the cytoskeleton in cultured cells (Nobes et al., 1995) and to affect cell migration and axon outgrowth in vivo (Luo et al., 1994). In addition,

constitutively active Rac1 rho-GTPase produced neuropathological changes in transgenic mice such as defects in axon outgrowth and dendritic spine morphogenesis in mouse Purkinje cells (Luo et al., 1996).

In the same manner, constitutive activation of a rho-GTPase resulting from the loss-of-function of oligophrenin 1 protein leads to a dysfunctioning of signal transduction pathways involved in cell migration and axon outgrowth during development of the nervous system. Mental retardation may be the clinical expression of such neuropathological changes.

The oligophrenin 1 gene would thus be involved in disorders due to an abnormal neurone migration. Such disorders include not only genetic disorders such as nonspecific X-linked mental retardation but also incurable cryptogenic epilepsies and neurodegenerative diseases, such as Alzheimer's disease and cognitive impairments related to aging.

The present invention also provides transgenic non-human animals or cells thereof.

Said transgenic animal can have a exogenous oligophrenin 1 protein of this invention due to the presence of a gene encoding and expressing that protein or part of that protein.

Transgenic animals are generally well known, as are their methods of production.

The present invention contemplates a non-human animal containing a oligophrenin 1 gene of the present invention integrated in the genome of the animal's somatic and germ cells, i.e., a transgenic animal, preferably transgenic mammals.

Animals containing a transgene encoding a oligophrenin 1 protein of the present invention are typically prepared using the standard transgenic technology described in Hogan et al. (1987) and Palmiter et al. (1986). Production of transgenic mammals is also possible using the homologous recombination transgenic systems described by Capecchi (1989). Preparation of transgenic mammals has also been described in WO94/21670.

One technique for transgenically altering a mammal is to microinject a rDNA into the male pronucleus of the fertilized mammalian egg to cause one or more copies of the rDNA to be retained in the cells of the developing mammal.

Alternative methods for producing a non-human mammal containing a rDNA of the present invention include infection of fertilized eggs, embryo-derived stem cells, totipotent embryonal carcinoma (Ec) cells, or early cleavage embryos with viral expression vectors containing the rDNA (Palmiter et al., (1986)).

A transgenic animal can also have a mutation in its own native oligophrenin 1 gene, thereby rendering the oligophrenin 1 protein non-functional (i.e., a "knockout" trangenic animal). Such an animal is useful as it presents the clinical conditions associated with the defects in the mutated oligophrenin 1 protein, and further can be a model for evaluation of candidate therapeutics that would treat subjects with defects in that protein.

More particularly, transgenic non-human animals or cells in culture, that overexpress oligophrenin 1 protein or preferably express a native oligophrenin 1 protein that has been rendered non-functional ("knock-out" transgenic animal) may be useful in a method for screening chemical entities or drugs likely to act on the signaling pathway to which the rho-GAP^{MRX} protein (oligophrenin) belongs.

Transgenic non-human animals or cells thereof that overexpress oligophrenin 1 protein refer to animals or cells thereof that express an exogenous oligophrenin 1 protein of the invention in addition not the native protein.

In one embodiment, the screening method of the invention comprises the steps of :

i) administering a drug to be tested to a transgenic non-human animal that overexpress oligophrenin 1 protein or preferably express a native oligophrenin 1 protein that has been rendered non-functional ; and

ii) observing clinical expression of neuronal changes *in vivo* and/or *in vitro* culturing nervous cells from said animal and observing the stimulation or recovery of axon outgrowth or morphogenesis.

In another embodiment, the screening method of the invention comprises the steps of :

i) contacting a drug to be tested with nervous cells or nervous tissue cultures that overexpress oligophrenin 1 protein or preferably express a native oligophrenin 1 protein that has been rendered non-functional ; and

ii) measuring the axon outgrowth.

In the above embodiment, said cells are either obtained from said transgenic animals or are established cell lines, such as neuroblastoma or primary cultures of neuronal cells, which have been transfected by a DNA construct, e.g by means of a viral vector, allowing the expression of exogenous oligophrenin 1 protein or rendering the native oligophrenin 1 non functional.

Drugs selected by the methods of screening as above-defined, and pharmaceutical compositions containing such a drug in association with a pharmaceutically acceptable carrier, are also encompassed by the present invention.

The ORF of the oligophrenin 1 gene as shown in SEQ ID n° 26 according to the invention encodes a protein of 802 amino acids with a relative molecular mass of 91 kD. Hydropathy analysis (Kyte and Doolittle, 1982) suggests that the oligophrenin 1 protein is hydrophilic. Based on consensus motifs in PROSITE database (Bairoch et al., 1997), many potential phosphorylation sites were predicted including a tyrosine kinase phosphorylation site at position 142. Comparison of the protein sequence with other sequences in the databases indicated that the oligophrenin 1 gene encodes a rho-GAP containing protein.

Sequence alignment shown in figure 3b illustrates the remarkable similarity between the predicted oligophrenin 1 domain and other members of the rho-GAP subfamily. This similarity extends over 180 residue region localised in the central part of the predicted protein and concerns the three

structurally conserved regions (SCRs) that are specific to the rho-GAP proteins.

Among rho-GAPs, the oligophrenin 1 protein showed the greatest similarity to the chicken Graf protein (Hildebrand et al., 1996). This similarity extends on both sides of the rho-GAP domains, but oligophrenin 1 does not contain the SH3 domain reported for the Graf protein. The rho-GAP activity of the oligophrenin 1 protein is consistent with the functional analysis of the chicken Graf protein, which has both part of the N-terminal and rho-GAP domains identified in the oligophrenin 1 protein. Graf protein has been shown to preferentially stimulate the GTPase activity of the GTP-binding proteins RhoA and Cdc42 (Hildebrand et al., 1996).

The C-terminal part of oligophrenin 1 protein does not match any known sequence, whereas the N-terminal domain of oligophrenin 1 protein is similar to a highly conserved protein, of unknown function, identified in *C. elegans*, mouse and human (Fig. 3c). This protein presents two isoforms identified as CELZK328 and CELT04C95 (Genbank), which correspond to two different ORF.

A further subject of the present invention is thus an isolated oligophrenin 1 polypeptide substantially comprising the aminoacid sequence of SEQ ID n° 27 or a homologous aminoacid sequence thereof.

The above term "substantially" is understood as meaning that said isolated oligophrenin 1 polypeptide exhibits the same biological and/or immunological properties, as the native oligophrenin 1 protein.

More particularly said aminoacid sequence may be SEQ ID n° 26, or a homologous aminoacid sequence thereof.

"A homologous aminoacid sequence" is understood as meaning a sequence which differs from the sequences to which it refers by mutation, insertion, deletion or substitution of one or more aminoacids, without inducing modification of biological and/or immunological properties. Said derivative

aminoacid sequence shows at least 60 % of homology, preferably 70 % of homology, preferably 80 % of homology with the oligophrenin 1 polypeptide of SEQ ID n° 26.

The "biological properties" of the polypeptides of the invention refer to the activity of the oligophrenin 1 protein, which enhances GTPase activity of small Ras-like GTPases and hence turns them off.

The "immunological properties" of the polypeptides of the invention refer to the ability of the polypeptides of the invention to induce an immunological response mediated by antibodies which recognize the oligophrenin 1 polypeptide of the invention.

The polypeptides according to the invention can be obtained by any of the standard methods of purification of soluble proteins, by peptide synthesis or by genetic engineering. Said techniques comprise the insertion of a nucleic acid sequence coding for a peptide of the invention into an expression vector, such as a plasmid, and the transformation of host cells with the expression vector, by any of the methods available to the skilled person, like for instance electroporation.

The present invention thus relates to vectors for cloning and/or expression comprising a nucleic acid sequence of the invention and to host cell transfected with these vectors. The expression vector according to the invention comprises a nucleic acid sequence encoding a polypeptide of the invention, said nucleic acid sequence being operably linked to elements allowing its expression. Said vector advantageously contains a promoter sequence, signals for initiation and termination of translation, as well as appropriate regions for regulation of translation. Its insertion into the host cell may be transient or stable. Said vector may also contain specific signals for secretion of the translated protein.

These various control signals are selected according to the host cell which may be inserted into vectors which self-replicate in the selected host cell, or into vectors which integrate the genome of said host.

Host cells may be prokaryotic or eukaryotic, including but not limiting to bacteria, yeasts, insect cells, mammalian cells, including cell lines which are commercially available.

A subject of the present invention is also a method for producing a recombining oligophrenin 1 polypeptide, wherein said host cell is transfected with said expression vector and is cultured in conditions allowing the expression of a polypeptide according to the invention.

The present invention also relates to monoclonal or polyclonal antibodies, or fragments thereof, or chimeric or immunoconjugate antibobies, which are capable of specifically recognizing a polypeptide according to the invention.

Polyclonal antibodies can be obtained from serum of an animal immunized against the oligophrenin 1, which can be produced by genetic engineering for example, as above described, according to standard methods well-known by one skilled in the art.

Monoclonal antibodies can be obtained according to the standard method of hybridoma culture (Kohler and Milstein, 1975).

The antibodies of the present invention can be chimeric antibodies, humanized antibodies, or antigen binding fragments Fab and F(ab')2. They can also be immunoconjugated or labelled antibodies.

Said antibodies are particularly useful for detecting or purifying a oligophrenin 1 polypeptide according to the invention in a biological sample.

They are more particularly useful for detecting an abnormal expression of the oligophrenin 1 protein in connection with neurological disorders as above described.

Another subject of the present invention is a pharmaceutical composition comprising a purified oligophrenin 1 polypeptide of the invention and/or a homologous polypeptide thereof, in association with a pharmaceutically acceptable carrier.

A further subject of the present invention is a pharmaceutical composition comprising a nucleic acid encoding said polypeptides and a pharmaceutically acceptable carrier. Said nucleic acid, preferably inserted in a vector, may be administered in a naked form or in association with transfection facilitating agents.

A further subject of the invention is a pharmaceutical composition comprising an anti-sense sequence capable of specifically hybridizing with a nucleic acid sequence encoding said polypeptides, in association with a pharmaceutically acceptable carrier.

A still further subject of the invention is a pharmaceutical composition comprising an antibody directed against said polypeptides, in association with a pharmaceutically acceptable carrier.

Preferably the present invention is directed to a pharmaceutical composition comprising a purified oligophrenin 1 polypeptide of the invention and/or a homologous polypeptide thereof, in association with a pharmaceutically acceptable carrier.

The term "homologous polypeptide", as active ingredient of a pharmaceutical composition, refers to a polypeptide with a homology of at least 40 %, preferably of at least 60 % in comparison to the oligophrenin 1 protein. Such homologous polypeptides include any known protein which exhibits a rho-GAP activity. Preferably said homologous polypeptide is for example the protein CELZK328 or CELT04C95.

The pharmaceutical compositions of the invention are useful for preventing and/or treating neurological disorders, wherein the oligophrenin 1 protein or a homologous protein thereof is implicated. As above underlined, the authors of the present invention have shown that defects in a Ras-like GTPase (rho-GAP) dependent signalling pathway are associated with cognitive impairment, resulting from misguided axon growth and/or defective cell migration. Consequently, the disorders which are more particularly aimed at are disorders of the central nervous system in connection with the axonal

development, more particularly a disorder associated with cognitive impairment. Such disorders include nonspecific X-linked mental retardation, as well as cryptogenic epilepsies or neurodegenerative diseases, such as Alzheimer's disease and cognitive impairments related to aging.

Another subject of the invention is the use of a purified oligophrenin 1 polypeptide of the invention and/or a homologous polypeptide thereof, in association with a pharmaceutically acceptable carrier for the manufacture of a medecine for preventing and/or treating neurological disorders, wherein the oligophrenin 1 protein or a homologous protein thereof is implicated.

The pharmaceutical compositions of the invention may be administered to a mammal, preferably to a human, in need of a such treatment, according to a dosage which may vary widely as a function of the age, weight and state of health of the patient, the nature and severity of the complaint and the route of administration.

The appropriate unit forms of administration comprise oral forms such as tablets, gelatin capsules, powders, granules and oral suspensions or solutions, sublingual and buccal administration forms, subcutaneous, intramuscular, intravenous, intranasal or intraoccular administration forms and rectal administration forms.

A further subject of the present invention is a method of preventing and/or treating neurological disorders resulting from defects in the oligophrenin 1 gene or in the oligophrenin 1 protein or in a homologous gene or protein thereof, which comprises administering to a subject in need of a such treatment an amount of a pharmaceutical composition as above defined effective to prevent and/or alleviate said neurological disorders.

The present invention is further illustrated by, but not limited to, the figures and the examples that follow.

LEGENDS TO FIGURES :

Figure 1. (**a**) Physical map of the Xq12 locus and genomic structure of the oligophrenin 1 gene spanning the X-chromosomal breakpoint. YAC, PAC and cosmid contigs are indicated with lines and STSs by vertical bars. The STS C16T3 was generated from the distal end of the 9 kb HindIII fragment containing the X-chromosomal breakpoint (represented by an arrow). The oligophrenin 1 gene spans at least 300 kb and it consists of 25 exons of which 23 are coding (white boxes). (**b**) Southern blot analysis of HindIII digested genomic DNAs from the patient exhibiting the (X;12) translocation and a normal female using the STS C16 T3 as probe. The junction fragment is indicated by JF.

Figure 2. Fetal and multiple adult tissue northern blots containing poly(A)$^+$ RNA (Clontech®) hybridized with C2 cDNA probe. A 7.5 kb transcript was observed after an overnight exposure at - 80°C. Hybridization of Northern blots with an actin probe was performed to assess differences between amounts of loaded poly(A)$^+$ RNA samples.

Figure 3. (**a**) Coding part of the cDNA and deduced amino acid sequences of the oligophrenin 1 gene. Nucleotides in bold correspond to exon-exon boundaries. The GAP domain is underlined, Primers used in RT-PCR to study the expression of the gene in the patient with the X;12 translocation are double underlined and the internal primer used for hybridization to ascertain PCR products is shown in dotted line. The deleted nucleotide at position 1578 is indicated in italic bold case. (**b**) Sequence alignment of the GAP domain present in oligophrenin 1 and different rhoGAP proteins reviewed in Aelst et al., 1997. The GAP domain contains three Structurally Conserved Regions (SCRs) (Boguski et al., 1993). CELZK328 corresponds to an ORF predicted from sequences of the C.*elegans* genome. (**c**). Sequence alignment showing

the high conservation of oligophrenin 1 N-terminal domain. CELTO4C95 corresponds to another C. elegans ORF, different from CELZK328. EST483210 is part of the mouse homologue oligophrenin 1 cDNA and EST387042 corresponds to a human EST that was localized on chromosome 11 by the Whitehead Institute.

Identical residues are indicated by black shading, similar residues by grey. The alignment was performed using Multalin (Corpet et al., 1988) and Boxshade softwares.

Figure 4. (**a**). Study of oligophrenin 1 transcript in the patient with the (X;12) translocation and in a normal female used as control (XX). Southern blot of PCR products amplified from total RNAs isolated from EBV transformed lymphoblastoid cell lines.

(**b**) Nucleotide sequences showing the one base pair deletion. Direct sequencing of PCR products corresponding to exon 19 of IV-3 proband.

(**c**). Segregation analysis of the mutated allele in XLMR family D. Open squares, unaffected males; closed squares, affected males; open circles, unaffected females; dotted symbols, phenotypically normal carrier females.

EXAMPLE 1 : Identification of the oligophrenin 1 gene :

1. Experimental procedures

Case report and family materials

Clinical data and diagnosis concerning the female patient with the t(X;12) translocation were previously described (Luo et al., 1996). Concerning the MRX families, a linkage study was reported in The European XLMR Consortium 's report (1997).

YAC and PAC clones

YAC clones of the Xq12 locus were obtained from the UK HGMP Resource Centre. PAC clones were obtained from the German resource center (RZPD). Primer sequences corresponding to STSs are available in Genome Data Base. STSC16T3 is a 189 bp fragment amplified with the primers: C16T3F (5' CACAGCAAGCAATAAGCACT 3') and C16T3R (5' TGTCTCCTGTGCTCTTTCCA 3'). Overlaps between clones and STS mapping were performed by a combination of STS/EST amplification and hybridization approaches.

cDNA isolation

Approximately 1 X 10^6 recombinant clones of a lgt10 human fetal brain cDNA library (Clontech) were plated and screened following standard techniques (Sambrook et al., 1989). Library screening was performed using as probe RT-PCR products generated with primers located within the predicted exons. Positive clones were plaque purified and their inserts were subcloned into bluescript vector and sequenced. 3' RACE PCR (Clontech kit) was used to obtain the 3' end of the cDNA. The full-length cDNA is a composite of 14 clones (only 4 clones are represented on figure 1a).

Genomic DNA sequencing of human cosmid clones:

Human cosmid clones were detected in the Imperial Cancer Research Fund (ICRF) flow-sorted human X chromosome library (Nizetic et al., 1991) using B2 cDNA as a probe. Cosmid 2C6, 4D2 and 35 shown in figure 1a are from a cosmid library corresponding to the YAC 4690. Exon-intron boundaries were identified through sequence comparisons between cDNA and genomic DNA clones. To generate genomic sequences, DNAs of cosmid clones were used as templates and primed with exonic oligonucleotides. ICRF coordinates of cosmid clones shown in figure 1 are as follow: cos12: ICRFc104J1515Q8, cos15: ICRFc104K1628Q8, cos7: ICRFc104P0212Q8, cos3: ICRFc104B1719Q8, cos11: ICRFc104F178Q8, cos5: ICRFc104B1515Q8.

2. Results :

As a first step in identifying a potential gene in Xq12 involved in MRX, Bienvenu *et al.* (1997) have reported a molecular cytogenetic investigation of an X;12 balanced translocation observed in a female affected with a mild mental retardation. The Xq breakpoint was localized within an ICRF 850 kb YAC clone y900H0493 (4690), positive for PGKP1 and DXS159 markers.

The authors of the present invention have used long-range restriction maps of the YAC clone 4690, FISH analyses and somatic hybrid cell lines containing the derivative chromosome 12 as their only human X-chromosomal component to regionally fine map the X chromosomal breakpoint. Figure 1a depicts the location of the translocation breakpoint on the normal X and summarizes YAC, PAC and cosmid contigs spanning the breakpoint. The probe, STSC16T3, which detects the junction fragment and localized the breakpoint to a 9 kb HindIII fragment (fig 1b) was isolated from the cosmid clone 4D2 (Fig 1a). Aberrant bands confirming the latter results

were also obtained by hybridization of the same probe to a Southern blot containing DNA from the patient digested with several other enzymes.

Sequencing of randomly subcloned HindIII fragments including the 9 kb fragment isolated from cosmid clones spanning the breakpoint, and searches for homology in data bases revealed sequence identities between the isolated sequences and those corresponding to the PAC clone 360E18 (Fig 1a) which were generated by the Sanger Centre (Cambridge,UK). Available sequences were then used for computational analyses and comparison with nucleotide and protein sequences. Some of the potential exons identified by GRAIL (Kel et al., 1993) and FEXH/HEXON (Lerman et al., 1987) programs showed a significant homology with the mouse EST 483210. The predicted polypeptide corresponding to this EST revealed a significant homology with the human EST 387042 localised on chromosome 11 and with the *C. elegans* ORF CELT04C95. Further investigations suggested that the *C. elegans* ORF is represented on the genomic sequences derived from the PAC clone by 8 different potential exons scattered over 130 kb (Fig. 1a). These predictions were confirmed by RT-PCR experiments using primers located within the potential exons and human fetal brain total RNA. Furthermore, hybridization of the RT-PCR products to a Northern blot containing polyA$^+$ RNA detected a 7.5 kb transcript most highly expressed in fetal brain (Fig. 2). Together, these results indicated the presence of a candidate gene located in the vicinity of the translocation breakpoint.

In order to identify the full length cDNA the authors of the present invention used a combination of fetal brain cDNA library screening, PCR and rapid amplifications of cDNA ends (RACE). This approach enabled to obtain a composite full length nucleotide sequence of the cDNA (Figure 3a).

To confirm that the identified gene is disrupted by the translocation breakpoint the structure of the gene including the exon-intron boundaries was determined through sequence comparisons between cDNA and cosmid genomic DNA clones isolated either from a cosmid library generated from YAC clone 4690 or the ICRF flow sorted X-specific cosmid

library (fig 1a). The physical mapping of the 25 exons allowed to demonstrate that the candidate MRX gene is transcribed from telomere to centromere and the translocation breakpoint maps within the second intron leading therefore to a displacement on the derivative chromosome 12 of the first two exons including the one containing the putative translation initiation codon. Confirmation of this latter result was obtained by FISH analysis using as probe a cosmid clone containing the first two exons of the gene (cos 12, fig 1a). This experiment showed that this cosmid maps exclusively on the derivative chromosome 12 (data not shown).

To investigate the gene expression and examine whether both alleles are inactive, RT-PCR experiments were performed on RNA isolated from EBV-transformed lymphoblastoid cell lines (LCL) of the patient and a female control using primers located in exons 1 and 2. These experiments failed to amplify a normal gene product on RNA of the patient (fig 4a). Consistently, the normal X chromosome was found to be late replicating, indicating preferential X-inactivation of the normal X chromosome (Bienvenu et al., 1997), and the MRX candidate gene was found to undergo X-inactivation.

EXAMPLE 2 : Identification of mutations in MRX families

1. Experimental procedure

Mutations analysis

Genomic DNA was extracted from EBV - transformed lymphoblastoid cell lines using standard protocols. The 23 coding exons and exon-intron boundaries were individually amplified with specific primers. In each amplification one primer was a 5' psoralen-modified primer (Fernandez et al., 1993). PCR products were checked on standard agarose gels prior to analysis by the DGGE technique. When an aberrant pattern of migration was observed, the corresponding PCR product was purified and directly sequenced

on both strands using the Dye Terminator Cycle Sequencing kit protocol (Applied Biosystems).

Figure 4(a) represents a Southern blot of PCR products amplified from total RNAs isolated from EBV transformed lymphoblastoid cell lines. RT-PCR were performed with (+) or without (-) reverse transcriptase and cDNAs were amplified for 40 cycles with primers located in exon 1 and 2 (figure 3, double underlined nucleotide sequences). After gel electrophoresis, the Southern blot was hybridized with an internal oligonucleotide (dotted line on figure 3). The 650 bp fragment corresponds to an amplification from the contaminating genomic DNA (intron 1 is 140 bp long). The lane labelled DNA, corresponds to the PCR product obtained from a female total genomic DNA. The negative control indicated by Ct(-) corresponds to a PCR experiment without template. RT-PCR amplification of the ubiquitously expressed transcript produced by the distal part of the dystrophin gene was used as internal standard.

The mutation to co-segregate with the mental retardation phenotype as shown on figure 4c was detected by denaturing gradient gel electrophoresis of PCR products corresponding to exon 19 of the oligophrenin 1 gene. Exon 19 was amplified by PCR with primers 19F (5' GTT AAT CTT GCC CCT TTT CT 3') and 19R (5' Psoralen- TA GGA AGA CAG GTA GTG AGA AT) yielding a 221 bp product. 10 µl of each amplified product was mixed with 10 µl of normal control PCR product. Heteroduplexes were generated by denaturing for 10 min, and subsequent reannealing for 45 min at 56°C. The samples were electrophoresed through a 25-65% denaturant 6% polyacrylamide gel for 7.5h at 60°C and 160V. The characteristic shifted profile displayed by the mutated allele allow an easy study of the familial segregation.

2. Results

In order to prove that the isolated gene is responsible for non-specific mental retardation, four unrelated probands from MRX families (The

European XLMR Consortium, 1997), previously mapped in genetic intervals which encompass the candidate gene, were analysed for the presence of point mutations. The strategy involved investigation by DGGE (Lerman et al., 1993) (denaturing gradient gel electrophoresis) of PCR products corresponding to all coding exons and sequencing of exons exhibiting abnormal migration profiles. PCR primers were designed not only to amplify individual exons but also sequences flanking the exons. DGGE analyses of amplified products, corresponding to exon 19, from the proband IV-3 of the family D (fig 4c) showed an abnormal shift in mobility. Compared with the normal product, sequence analysis showed that the aberrant product contained a one base pair deletion of the nucleotide 1578 (fig 4b); the resulting frameshift mutation was predicted to cause premature translation termination four codons downstream of the mutation. Cosegregation of the mutation with the disease which was confirmed in the large family using the DGGE technique (fig 2b) and the absence of this mutation in 100 control individuals indicate that the deletion does indeed cause the mutant phenotype.

EXAMPLE 3 : Expression of the oligophrenin 1 gene

C2 cDNA clone was isolated from a fetal brain cDNA library. C2 DNA probe was used to hybridize poly(A)$^+$ RNA (Clontech©) contained in fetal and multiple adult tissues. A 7.5 kb transcript was observed after an overnight exposure at - 80°C. Hybridization of Northern blots with an actin probe was performed to assess differences between amounts of loaded poly(A)$^+$ RNA samples.

As shown in figure 2, the oligophrenin 1 transcript was most abundant in RNA from fetal brain. A lower level of expression was also detected in several other tissues including adult brain. To further investigate the distribution of the transcript during development, *in situ* hybridization was used to examine the expression of the mouse homologous gene in embryonic days (E) 10.5, E12.5, E14, E18 and in postnatal day 1 of mouse embryos and

postnatal tissues. In addition to a low expression in all tissues with no significant differences, it was found that the gene is expressed at a higher level in all parts of the developing neuroepithelium of the neural tube. During later differentiation stages and in the mature brain a significant level of expression is visible in different structures of the brain with no striking distribution of the mRNA expression.

Several lines of evidence show that defects in oligophrenin 1 are responsible for X-linked non-specific mental retardation. First, the oligophrenin 1 gene maps to a potential mental retardation genetic locus in Xq12 identified by linkage analyses (Lubs et al., 1996, The European XLMR Consortium, 1997). Second, literature reports (Davies, 1997) concerning two patients with complete androgen insensitivity syndrome (CAIS) and mental retardation showed the presence of deletions which include several markers both proximal and distal to the AR gene and extend to DXS905 and DXS908. The above mapping data showed that these markers are located within the second and fifth intron of the oligophrenin 1 gene (Fig. 1a), demonstrating therefore that most of the exons are deleted in these two patients with mental retardation. In contrast, deletions in two CAIS patients without mental retardation do not extend to the oligophrenin 1 gene as deletions are limited to the androgen receptor gene itself (Davies, 1997). Third, investigation of this gene in the female patient with mental retardation and a t(X;12) demonstrated an absence of expression of both alleles resulting from the disruption of one allele by the translocation breakpoint and a preferential inactivation of the second allele. Fourth, a one base pair deletion within the rho-GAP domain, predicted to result in a severe abbreviated and nonfunctional oligophrenin 1 protein, cosegregates with a recessive mental retardation phenotype in a large affected family mapped to the pericentromeric region. Finally, the oligophrenin 1 mRNA is highly expressed in fetal brain, a finding consistent with the disease phenotype.

REFERENCES

1. Aelst, L.V. & D'Souza-Schorey, C. Rho GTPases and signaling networks. *Genes & Dev.* **11**, 2295-2322 (1997).
2. Bairoch, A., Bucher, P. & Hofmann, K. The PROSITE database, its status in 1997. *Nucleic Acids Res* **25**, 217-21 (1997).
3. Bienvenu, T., *et al.* Mapping of the X-breakpoint involved in a balanced X;12 translocation in a female with mild mental retardation. *Eur. J. of Hum Genet* **5**, 105-109 (1997).
4. Boguski, M.S. & McCormick, F. Proteins regulating Ras and its relatives. *Nature* **366**, 643-54 (1993).
5. Capecchi et al., *Science* **244**, 1288-1292 (1989).
6. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**, 10881-90 (1988).
7. Davies, H.R., *et al.* Androgen insensitivity with mental retardation: a contiguous gene syndrome ? *J Med Genet* **34**, 158-60 (1997).
8. Fernandez, E., *et al.* Use of chemical clamps in denaturing gradient gel electrophoresis: application in the detection of the most frequent Mediterranean beta-thalassemic mutations. *PCR Methods Appl* **3**, 122-4 (1993).
9. Gecz, J., *et al.* Cloning and expression of the murine homologue of a putative human X-linked nuclear protein gene closely linked to PGK1 in Xq13.3. *Hum Mol Genet* **3**, 39-44 (1994).
10. Gedeon, A.K., Donnelly, A.J., Mulley, J.C., Kerr, B. & Turner, G. How many X-linked genes for non-specific mental retardation (MRX) are there? [letter]. *Am J Med Genet* **64**, 158-62 (1996).
11. Gecz J. et al., Identification of the gene FMR2, associated with FRAXE mental retardation. *Nature Genet* **13**, 105-108 (1996).
12. Gibbons RJ et al., Syndrome mental retardation due to mutations in a regulator of gene expression. *Hum Mol Genet* **4**, 1705-1709 (1995).

13. Herbst, D.S. & Miller, J.R. Nonspecific X-linked mental retardation II: the frequency in British Columbia. *Am J Med Genet* **7**, 461-9 (1980).

14. Hildebrand, J.D., Taylor, J.M. & Parsons, J.T. An SH3 domain-containing GTPase-activating protein for Rho and Cdc42 associates with focal adhesion kinase. *Mol Cell Biol* **16**, 3169-78 (1996).

15. Hogan et al., Cold Spring Harbor, NY *Manipulating the mouse embryo ; a laboratory manual* (1987).

16. Kel, A.E., *et al.* SITEVIDEO: a computer system for functional site analysis and recognition. Investigation of the human splice sites. *Comput Appl Biosci* **9**, 617-27 (1993).

17. Knight SJI et al., Trinucleotide repeat amplification and hypermethylation of CpG islands in FRAXE mental retardation. *Cell* **74**, 127-134 (1993).

18. Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-92 (1986).

19. Kyte, J. & Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32 (1982).

20. Lamarche, N. & Hall, A. GAPs for rho-related GTPases. *Trends Genet* **10**, 436-40 (1994).

21. Lerman, L.S. & Silverstein, K. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. *Methods Enzymol* **155**, 482-501 (1987).

22. Lubs, H.A., *et al.* XLMR genes: update 1996. *Am J Med Genet* **64**, 147-57 (1996).

23. Luo, L., Liao, Y.J., Jan, L.Y. & Jan, Y.N. Distinct morphogenetic functions of similar small GTPases: Drosophila Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes Dev* **8**, 1787-802 (1994).

24. Luo, L., *et al.* Differential effects of the Rac GTPase on Purkinje cell axons and dendritic trunks and spines. *Nature* **379**, 837-40 (1996).

25. Mermod, N., O'Neill, E.A., Kelly, T.J. & Tjian, R. The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell* **58**, 741-53 (1989).

26. Morrissey, J., *et al.* A serine/proline-rich protein is fused to HRX in t(4; 11) acute leukemias. *Blood* **81**, 1124-31 (1993).

27. Nobes, C.D. & Hall, A. Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* **81**, 53-62 (1995).

28. Nizetic, D., *et al.* Construction, arraying, and high-density screening of large insert libraries of human chromosomes X and 21: their potential use as reference libraries. *Proc Natl Acad Sci U S A* **88**, 3233-7 (1991).

29. Palmiter et al., *Ann Rev Genet* **20**, 465-499 (1986).

30. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning: A laboratory manual,* Cold Spring Harbor Laboratory Press (1989).

31. Solovyev, V.V., Salamov, A.A. & Lawrence, C.B. Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Nucleic Acids Res* **22**, 5156-63 (1994).

32. The European XLMR Consortium. X-linked non specific mental retardation (MRX): linkage studies in 25 unrelated families. *8th International Workshop on the Fragile X and X-Linked Mental Retardation. Picton; Ontario, Canada, 17-22 August 1997.* (1997).

33. Verkerk AJM et al., Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **74**:905-914 (1991).

34. Vits L. et al., MASA syndrome is due to mutations in the neural cell adhesion gene LICAM. *Nature Genet* **7**, 408-413 (1994).

35. Zipkin, I.D., Kindt, R.M. & Kenyon, C.J. Role of a new rho family member in cell migration and axon guidance in C.elegans. *Cell* **90**, 883-894 (1997).

MODIFIED CLAIMS

1. Nucleic acid, **the sequence of which** is selected from the group consisting of sequences SEQ ID n° 1 to SEQ ID n° 25, and a homologous nucleic acid sequence thereof.

2. Nucleic acid, **the sequence of which** is selected from the group consisting of exon sequences as identified in table 2, and a homologous nucleic acid sequence thereof.

3. Nucleic acid comprising the sequence as shown in SEQ ID n° 26, a homologous sequence thereof, or a sequence identical to SEQ ID n° 26, except for a one base deletion of the nucleotide 1578 as shown in SEQ ID N° 26.

4. Isolated oligophrenin 1 polypeptide substantially comprising the aminoacid sequence of SEQ ID n° 27, or a homologous amino acid sequence thereof.

5. Vector for cloning and/or expression comprising a nucleic acid sequence of any of claims 1 to 3.

6. Host cell transfected with a vector according to claim 5.

7. Nucleic acid, **the sequence of which has at least 15 bases** and specifically hybridizes with a nucleic acid sequence according to any **of claims 1 to 2, under stringent conditions.**

8. Nucleic acid, **the sequence of which has at least 15 bases** and specifically hybridizes with a nucleic acid sequence according to **claim 3 under stringent conditions.**

9. Nucleic acid of claim 7 or 8, **the sequence of which** is selected from the group consisting of the sequences identified in table 3 or the complementary sequences thereof, said sequences identified in table 3 consisting of:

- nucleotide n° 727 to 746 of SEQ ID n° 2

- nucleotide n° 958 to 977 of SEQ ID n° 2
- nucleotide n° 375 to 394 of SEQ ID n° 3
- nucleotide n° 504 to 523 of SEQ ID n° 3
- nucleotide n° 418 to 437 of SEQ ID n° 4
- nucleotide n° 551 to 570 of SEQ ID n° 4
- nucleotide n° 423 to 445 of SEQ ID n° 5
- nucleotide n° 553 to 574 of SEQ ID n° 5
- nucleotide n° 388 to 407 of SEQ ID n° 6
- nucleotide n° 540 to 559 of SEQ ID n° 6
- nucleotide n° 436 to 458 of SEQ ID n° 7
- nucleotide n° 584 to 603 of SEQ ID n° 7
- nucleotide n° 219 to 239 of SEQ ID n° 8
- nucleotide n° 363 to 381 of SEQ ID n° 8
- nucleotide n° 108 to 128 of SEQ ID n° 9
- nucleotide n° 336 to 355 of SEQ ID n° 9
- nucleotide n° 361 to 380 of SEQ ID n° 10
- nucleotide n° 492 to 511 of SEQ ID n° 10
- nucleotide n° 81 to 100 of SEQ ID n° 11
- nucleotide n° 223 to 242 of SEQ ID n° 11
- nucleotide n° 188 to 207 of SEQ ID n° 12
- nucleotide n° 300 to 319 of SEQ ID n° 12
- nucleotide n° 166 to 189 of SEQ ID n° 13
- nucleotide n° 259 to 278 of SEQ ID n° 13
- nucleotide n° 133 to 152 of SEQ ID n° 14
- nucleotide n° 250 to 269 of SEQ ID n° 14
- nucleotide n° 151 to 170 of SEQ ID n° 15
- nucleotide n° 293 to 315 of SEQ ID n° 15
- nucleotide n° 221 to 244 of SEQ ID n° 16
- nucleotide n° 363 to 382 of SEQ ID n° 16
- nucleotide n° 305 to 324 of SEQ ID n° 17
- nucleotide n° 438 to 457 of SEQ ID n° 17
- nucleotide n° 25 to 44 of SEQ ID n° 18
- nucleotide n° 218 to 237 of SEQ ID n° 18
- nucleotide n° 51 to 70 of SEQ ID n° 19

- **nucleotide n° 252 to 271 of SEQ ID n° 19**
- **nucleotide n° 206 to 225 of SEQ ID n° 20**
- **nucleotide n° 383 to 402 of SEQ ID n° 20**
- **nucleotide n° 151 to 170 of SEQ ID n° 21**
- **nucleotide n° 511 to 530 of SEQ ID n° 21**
- **nucleotide n° 294 to 313 of SEQ ID n° 22**
- **nucleotide n° 496 to 515 of SEQ ID n° 22**
- **nucleotide n° 179 to 198 of SEQ ID n° 23**
- **nucleotide n° 271 to 291 of SEQ ID n° 23**
- **nucleotide n° 65 to 84 of SEQ ID n° 24**
- **nucleotide n° 165 to 184 of SEQ ID n° 24**
- **nucleotide n° 152 to 172 of SEQ ID n° 26**
- **nucleotide n° 586 to 606 of SEQ ID n° 26**
- **nucleotide n° 641 to 663 of SEQ ID n° 26**

10. Method for producing a **recombinant** oligophrenin 1 polypeptide, wherein a host cell of claim **6** is cultured in conditions allowing the expression of a polypeptide according to claim **4**.

11. Monoclonal or polyclonal antibodies, or fragments thereof, chimeric or immunoconjugate antibobies, which are capable of specifically recognizing a polypeptide according to claim **4**.

12. Use of the antibodies of claim 11 for detecting or purifiyng a polypeptide according to claim **4** in a biological sample.

13. Use of a nucleic acid according to any of claims 1, 2, 3, **7**, **8** and 9, for detecting an abnormality in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene.

14. Method of *in vitro* diagnosis of a neurological disorder associated with an abnormality in the oligophrenin 1 gene or in the transcripts of the oligophrenin 1 gene, wherein one or more mutation(s) is detected **in any of the sequences of claims 1 to 3.**

15. Method according to claim 14 wherein said mutation is a one base deletion of the nucleotide 1578 as shown in SEQ ID N° 26.

16. Method of *in vitro* diagnosis according to any of claims 14 or 15 comprising the steps of :

- contacting a biological sample containing DNA with specific oligonucleotides **having a sequence as defined in claim 7**, permitting the amplification of all or part of the oligophrenin 1 gene, the DNA contained in the sample having being rendered accessible, where appropriate, to hybridization, and under conditions permitting a hybridization of the oligonucleotides with the DNA contained in the biological sample ;
- amplifying said DNA ;
- detecting the amplification products ;
- comparing the amplified products as obtained to the amplified products obtained with a normal control biological sample, and thereby detecting a possible abnormality in the oligophrenin 1 gene.

17. Method of *in vitro* diagnosis according to any of claims 14 or 15 comprising the steps of :

- producing cDNA from mRNA contained in a biological sample ;
- contacting said cDNA with specific oligonucleotides **having a sequence as defined in claim 8,** permitting the amplification of all or part of the transcript of the oligophrenin 1 gene, under conditions permitting a hybridization of the primers with said cDNA ;
- amplifying said cDNA ;
- detecting the amplification products ;
- comparing the amplified products as obtained to the amplified products obtained with a normal control biological sample, and thereby detecting a possible abnormality in the transcript of the oligophrenin 1 gene.

18. Pharmaceutical composition comprising a purified oligophrenin 1 polypeptide of **claim 4** and/or a homologous polypeptide thereof, or an isolated

nucleic acid sequence encoding said polypeptide in association with a pharmaceutically acceptable carrier

19. Pharmaceutical composition comprising an anti-sense sequence **as defined in claim 7 or 8** in association with a pharmaceutically acceptable carrier.

20. Pharmaceutical composition comprising an **antibody according to claim 11.**

21. Transgenic **non-human** mammal expressing an exogenous oligophrenin 1 protein as defined in claim 4, **or being modified so as to overexpress a native oligophrenin 1 protein as defined in claim 4, or so as to express a non-functional oligophrenin 1 protein as defined in claim 4.**

22. Method for screening drugs likely to act on the signaling pathway to which the oligophrenin 1 protein belongs, wherein said drugs are tested on transgenic non-human mammals, or cells in culture, that overexpress oligophrenin 1 protein **as defined in claim 4** or express a native oligophrenin 1 protein **as defined in claim 4** that has been rendered non-functional.

23. Drug selected by the method of claim 22.

24. Pharmaceutical composition containing a drug of claim 23 in association with a pharmaceutically acceptable carrier.

25. Method of preventing and/or treating neurological disorders resulting from defects in the oligophrenin 1 gene or in the oligophrenin 1 protein or in a homologous gene or protein thereof, which comprises administering to a subject in need of a such treatment an amount of a pharmaceutical composition of claim 18 or 24 effective to prevent and/or alleviate said neurological disorders.

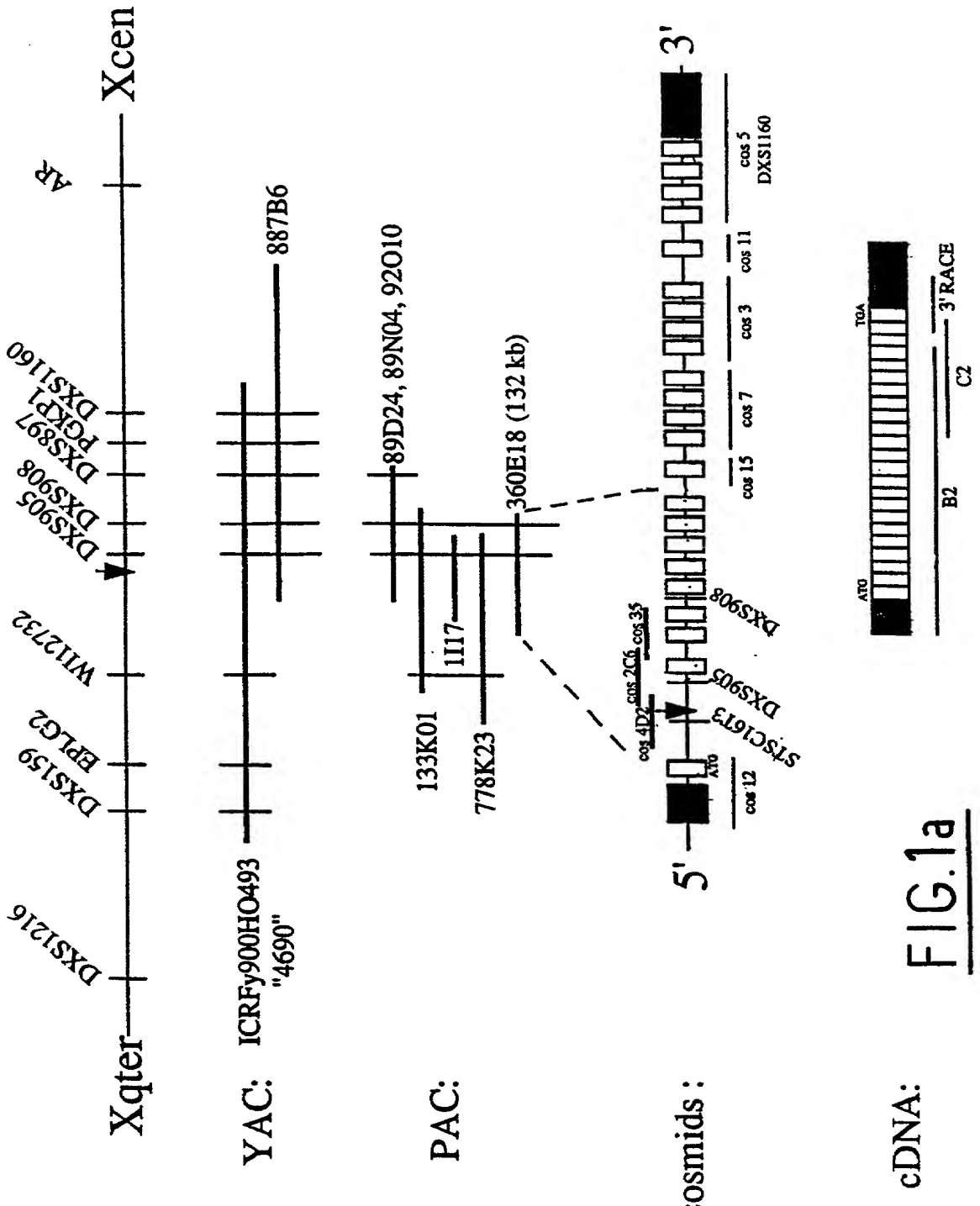
Xqter
Xcen
DXS1216
DXS159
EPLG2
WI12732
DXS905
DXS908
PGKP1
DXS897
DXS1160
AR
YAC:
ICRFy900HO493
"4690"
887B6
PAC:
89D24, 89N04, 92O10
360E18 (132 kb)
133K01
1117
778K23
cosmids :
5'
3'
ATG
cos 12
cos 4D2
cos 2C6
cos 35
DXS905
STSC16173
DXS908
cos 15
cos 7
cos 3
cos 11
cos 5
DXS1160
cDNA:
ATG
TGA
B2
C2
3'RACE

FIG.1a

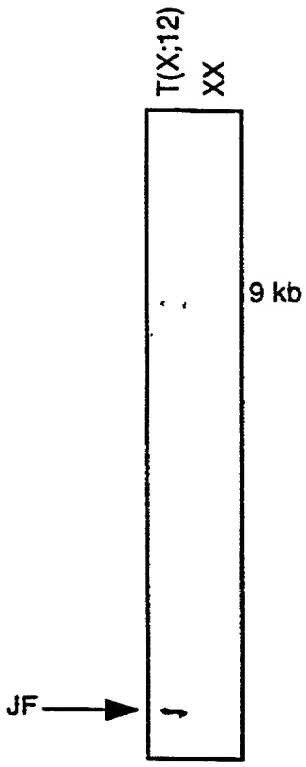
T(X;12)
XX
9 kb
JF

FIG. 1b

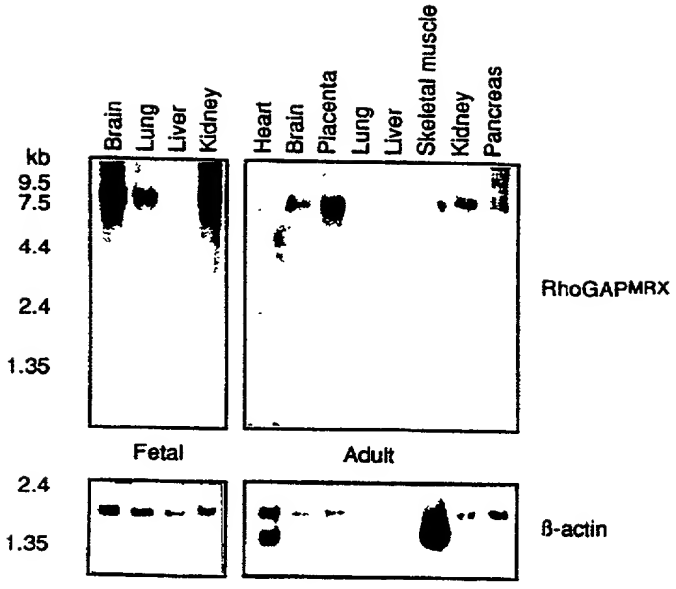
Brain
Lung
Liver
Kidney
Heart
Brain
Placenta
Lung
Liver
Skeletal muscle
Kidney
Pancreas
kb
9.5
7.5
4.4
2.4
1.35
RhoGAPMRX
Fetal
Adult
2.4
1.35
ß-actin

FIG. 2

```
      tgtgcgtcgcgctctcgcctcctcttcccgctgcagtgtctatggagcgaggctacgtttcattgccgccctggcttaacccttc  -553
cggcgcctaaaaggacggccggccggccggtcccttgcaccaggaagaagtcttagcagccagcgggccctggtcaggaaactctaaggtac  -461
aaggaaaacagttgaggaaggagccagagcgctccggtttggtcctcgggcttcgctggggcggggcgcaggcgttggctttaagaaagggg  -369
aggggacagtgcaatccgggttgcccgcggattcggccaaggaatcttccgctcgctccggagcgaggagcctgtaaagaggctgttcccag  -277
ctccagctctaacctcgcctacaccttgggcgggcccaatgtcacgtttgcaattgctcaggaaggatccggcccgtctccggaggcaagtc  -185
gggctgcggtttttgctgcttatctgggaaggcgatgcctaagggacatgctgcttgctaggcagcaccctgccgggatccgactgcgatag   -93
ttagctctccctggccctgaagccatcgccggggcgcctgttctctgtccggaccagccagcgctcctcaggagtctcactgaaacagaacc    -1
ATG GGT CAT CCC CCG CTG GAG TTC AGC GAC TGC TAC CTG GAC AGC CCC GAT TTC CGC GAG AGG CTC AAG      69
 M   G   H   P   P   L   E   F   S   D   C   Y   L   D   S   P   D   F   R   E   R   L   K   23
TGT TAT GAG CAG GAA CTG GAG AGG ACC AAC AAA TTC ATC AAA GAC GTA ATC AAA GAC GGC AAC GCG CTT    138
 C   Y   E   Q   E   L   E   R   T   N   K   F   I   K   D   V   I   K   D   G   N   A   L   46
ATC AGC GCT ATG AGA AAT TAT TCT TCT GCT GTT CAG AAA TTT TCC CAG ACG CTG CAG TCA TTT CAG TTT    207
 I   S   A   M   R   N   Y   S   S   A   V   Q   K   F   S   Q   T   L   Q   S   F   Q   F   69
GAT TTC ATT GGA GAC ACT CTG ACT GAT GAT GAA ATT AAC ATC GCT GAA TCC TTC AAG GAA TTT GCT GAA    276
 D   F   I   G   D   T   L   T   D   D   E   I   N   I   A   E   S   F   K   E   F   A   E   92
TTG CTC AAC GAG GTA GAA AAT GAG AGG ATG ATG ATG GTA CAC AAT GCT AGT GAT TTG CTG ATT AAA CCC    345
 L   L   N   E   V   E   N   E   R   M   M   M   V   H   N   A   S   D   L   L   I   K   P  115
TTG GAA AAT TTC CGG AAG GAA CAA ATA GGC TTC ACC AAG GAG CGG AAA AAG AAA TTT GAA AAG GAT GGT    414
 L   E   N   F   R   K   E   Q   I   G   F   T   K   E   R   K   K   K   F   E   K   D   G  138
GAG AGG TTT TAT TCT TTA CTG GAT CGG CAC TTA CAC CTG TCT TCA AAA AAG AAA GAA TCT CAG TTA CAA    483
 E   R   F   Y   S   L   L   D   R   H   L   H   L   S   S   K   K   K   E   S   Q   L   Q  161
GAG GCA GAC CTA CAG GTG GAC AAG GAG AGG CAC AAT TTT TTC GAG TCC TCT CTT GAT TAT GTT TAT CAA    552
 E   A   D   L   Q   V   D   K   E   R   H   N   F   F   E   S   S   L   D   Y   V   Y   Q  184
ATC CAG GAA GTT CAG GAG TCC AAG AAG TTC AAT ATT GTG GAG CCT GTC TTG GCC TTT CTT CAT AGT CTG    621
 I   Q   E   V   Q   E   S   K   K   F   N   I   V   E   P   V   L   A   F   L   H   S   L  207
TTC ATT TCT AAC AGC CTG ACT GTG GAG CTC ACA CAG GAT TTC CTC CCA TAC AAA CAA CAG CTC CAA CTC    690
 F   I   S   N   S   L   T   V   E   L   T   Q   D   F   L   P   Y   K   Q   Q   L   Q   L  230
AGT TTA CAG AAT ACA AGA AAT CAT TTC TCC AGT ACC CGG GAA GAG ATG GAA GAA CTT AAG AAA AGG ATG    759
 S   L   Q   N   T   R   N   H   F   S   S   T   R   E   E   M   E   E   L   K   K   R   M  253
AAA GAA GCT CCC CAG ACA TGC AAA CTT CCA GGA CAG CCA ACT ATT GAA GGC TAT CTC TAT ACA CAA GAG    828
 K   E   A   P   Q   T   C   K   L   P   G   Q   P   T   I   E   G   Y   L   Y   T   Q   E  276
AAA TGG GCT TTA GGA ATA TCC TGG GTG AAA TAC TAT TGC CAG TAT GAG AAA GAG ACC AAA ACA CTG ACC    897
 K   W   A   L   G   I   S   W   V   K   Y   Y   C   Q   Y   E   K   E   T   K   T   L   T  299
ATG ACG CCT ATG GAG CAG AAG CCA GGT GCT AAG CAG GGG CCC TTG GAC TTA ACA CTG AAG TAC TGT GTG    966
 M   T   P   M   E   Q   K   P   G   A   K   Q   G   P   L   D   L   T   L   K   Y   C   V  322
AGA AGG AAG ACG GAG TCT ATC GAC AAG AGG TTC TGT TTT GAC ATA GAA ACT AAT GAA AGG CCA GGA ACC   1035
 R   R   K   T   E   S   I   D   K   R   F   C   F   D   I   E   T   N   E   R   P   G   T  345
ATC ACT CTG CAG GCC CTT TCA GAA GCT AAC AGA AGG CTA TGG ATG GAA GCC ATG GAT GGG AAA GAA CCT   1104
 I   T   L   Q   A   L   S   E   A   N   R   R   L   W   M   E   A   M   D   G   K   E   P  368
```

FIG.3a

```
ATC TAC CAC AGC CCT ATA ACA AAA CAG CAA GAA ATG GAG CTA AAT GAA GTG GGC TTC AAG TTT GTC AGG   1173
 I   Y   H   S   P   I   T   K   Q   Q   E   M   E   L   N   E   V   G   F   K   F   V   R  391
AAG TGC ATC AAT ATT ATT GAG ACC AAA GGG ATC AAG ACA GAA GGG TTG TAC CGC ACT GTG GGC AGC AAT   1242
 K   C   I   N   I   I   E   T   K   G   I   K   T   E   G   L   Y   R   T   V   G   S   N  414
ATT CAG GTT CAG AAG CTG CTG AAT GCC TTT TTT GAT CCT AAA TGC CCA GGA GAT GTT GAT TTT CAT AAT   1311
 I   Q   V   Q   K   L   L   N   A   F   F   D   P   K   C   P   G   D   V   D   F   H   N  437
AGT GAC TGG GAC ATT AAG ACA ATC ACC AGC TCC TTG AAA TTC TAC CTC AGG AAT CTT TCT GAA CCT GTC   1380
 S   D   W   D   I   K   T   I   T   S   S   L   K   F   Y   L   R   N   L   S   E   P   V  460
ATG ACC TAT AGA CTT CAC AAA GAG CTG GTC TCT GCT GCC AAG TCT GAC AAC CTG GAT TAC CGC CTA GGA   1449
 M   T   Y   R   L   H   K   E   L   V   S   A   A   K   S   D   N   L   D   Y   R   L   G  483
GCT ATT CAC TCC CTG GTA TAT AAG CTA CCA GAA AAG AAC CGA GAG ATG CTG GAA CTT CTG ATA AGA CAC   1518
 A   I   H   S   L   V   Y   K   L   P   E   K   N   R   E   M   L   E   L   L   I   R   H  506
TTG GTC AAT GTG TGT GAG CAC AGC AAA GAG AAT CTT ATG ACC CCC TCC AAC ATG GGA GTA ATC TTT GGG   1587
 L   V   N   V   C   E   H   S   K   E   N   L   M   T   P   S   N   M   G   V   I   F   G  529
CCC ACC CTG ATG AGA GCT CAG GAG GAC ACT GTG GCC GCC ATG ATG AAC ATC AAA TTC CAG AAC ATA GTG   1656
 P   T   L   M   R   A   Q   E   D   T   V   A   A   M   M   N   I   K   F   Q   N   I   V  552
GTG GAA ATA CTA ATC GAG CAC TTT GGC AAG ATC TAT TTA GGT CCA CCT GAG GAA AGC GCT GCA CCG CCA   1725
 V   E   I   L   I   E   H   F   G   K   I   Y   L   G   P   P   E   E   S   A   A   P   P  575
GTG CCT CCG CCT CGG GTG ACA GCA AGA AGG CAC AAA CCA ATC ACG ATT TCA AAG CGC TTG CTG CGA GAA   1794
 V   P   P   P   R   V   T   A   R   R   H   K   P   I   T   I   S   K   R   L   L   R   E  598
AGG ACG GTT TTC TAT ACT TCT TCC CTG GAT GAA AGC GAA GAT GAA ATC CAA CAT CAA ACA CCG AAT GGT   1863
 R   T   V   F   Y   T   S   S   L   D   E   S   E   D   E   I   Q   H   Q   T   P   N   G  621
ACT ATC ACC AGC AGC ATA GAA CCC CCC AAG CCA CCA CAA CAC CCC AAA CTA CCT ATT CAG AGG AGT GGG   1932
 T   I   T   S   S   I   E   P   P   K   P   P   Q   H   P   K   L   P   I   Q   R   S   G  644
GAA ACT GAT CCT GGG AGG AAG TCC CCA AGC AGG CCT ATT TTG GAT GGC AAG TTG GAG CCC TGC CCA GAG   2001
 E   T   D   P   G   R   K   S   P   S   R   P   I   L   D   G   K   L   E   P   C   P   E  667
GTG GAC GTG GGG AAG TTG GTG TCT AGG CTG CAG GAT GGA GGG ACC AAG ATC ACC CCA AAG GCC ACC AAT   2070
 V   D   V   G   K   L   V   S   R   L   Q   D   G   G   T   K   I   T   P   K   A   T   N  690
GGA CCC ATG CCA GGC TCT GGG CCC ACC AAG ACC CCC TCT TTC CAC ATA AAG AGA CCA GCT CCC CGG CCC   2139
 G   P   M   P   G   S   G   P   T   K   T   P   S   F   H   I   K   R   P   A   P   R   P  713
CTG GCC CAC CAC AAG GAG GGG GAT GCT GAC AGT TTC AGC AAA GTG CGG CCT CCA GGA GAA AAG CCA ACC   2208
 L   A   H   H   K   E   G   D   A   D   S   F   S   K   V   R   P   P   G   E   K   P   T  736
ATC ATC CGC CCC CCA GTG AGG CCC CCA GAT CCT CCC TGC CGG GCA GCT ACT CCC CAA AAG CCA GAA CCA   2277
 I   I   R   P   P   V   R   P   P   D   P   P   C   R   A   A   T   P   Q   K   P   E   P  759
AAG CCA GAT ATT GTG GCT GGC AAT GCG GGG GAA ATC ACA TCA TCT GTG GTG GCT TCC AGG ACC AGG TTT   2346
 K   P   D   I   V   A   G   N   A   G   E   I   T   S   S   V   V   A   S   R   T   R   F  782
TTT GAA ACA GCT TCC CGG AAA ACA GGA AGT TCT CAA GGC AGA CTT CCT GGA GAT GAA AGT TGA ggctaca   2416
 F   E   T   A   S   R   K   T   G   S   S   Q   G   R   L   P   G   D   E   S   *          802
ggttttaaaagccttggcctcagaggacccttccaggttctgaaag...                                          2463
```

FIG. 3a cont.

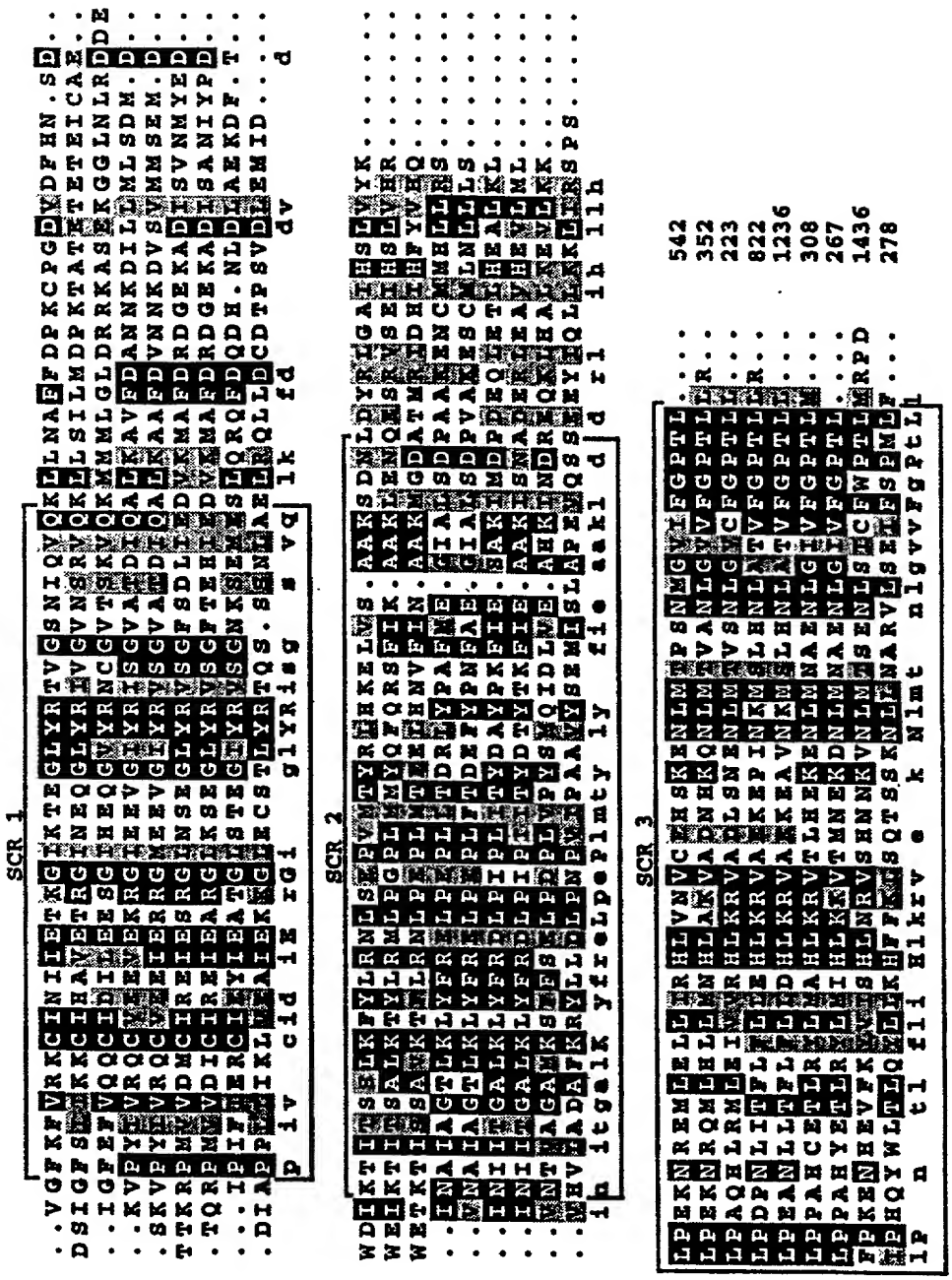
SCR 1
SCR 2
SCR 3
rhoGAPMRX 392
graf 197
CELZK328 69
Abr 672
Bcr 1086
N-chimaerin 156
β-chimaerin 117
p190 1286
p85α 128
consensus
rhoGAPMRX
graf
CELZK328
Abr
Bcr
N-chimaerin
β-chimaerin
p190
p85α
consensus
rhoGAPMRX 542
graf 352
CELZK328 223
Abr 822
Bcr 1236
N-chimaerin 308
β-chimaerin 267
p190 1436
p85α 278
consensus

FIG. 3b

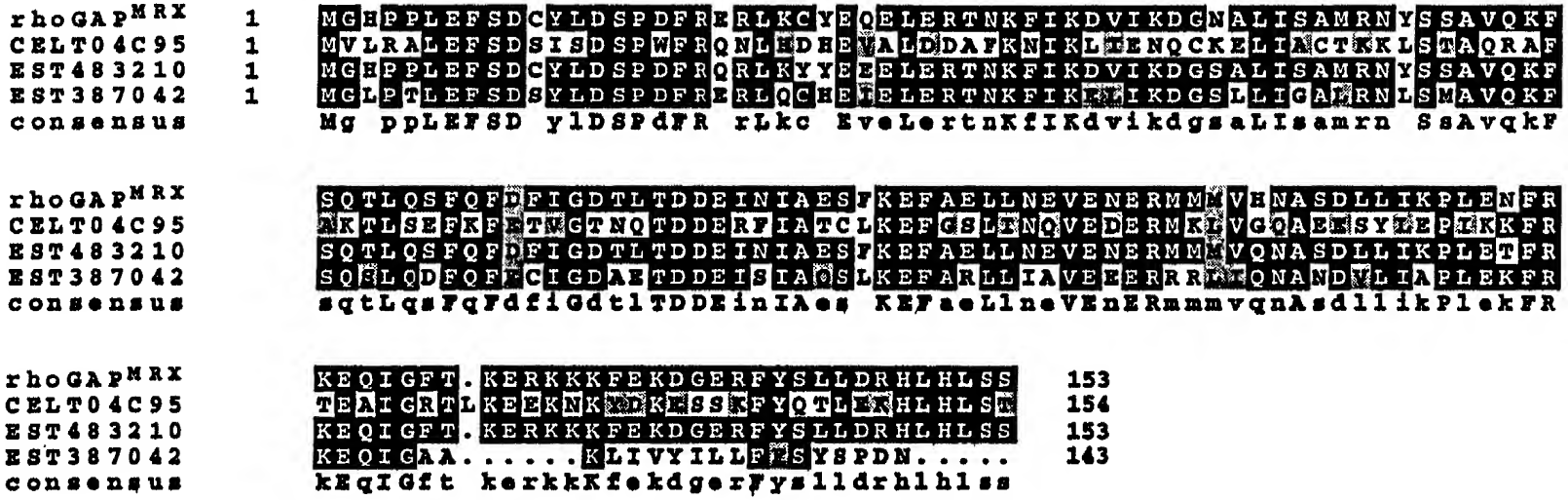
rhoGAP^MRX 1 MGHPPLEFSDCYLDSPDFRERLKCYEQELERTNKFIKDVIKDGNALISAMRNYSSAVQKF
CELT04C95 1 MVLRALEFSDSISDSPWFRQNLHDHEVALDDAFKNIKLTENQCKELIACTKKLSTAQRAF
EST483210 1 MGHPPLEFSDCYLDSPDFRQRLKYYEEELERTNKFIKDVIKDGSALISAMRNYSSAVQKF
EST387042 1 MGLPTLEFSDSYLDSPDFRERLQCHETELERTNKFIKELIKDGSLLIGAMRNLSMAVQKF
consensus Mg ppLEFSD ylDSPdFR rLkc EveLertnKfIKdvikdgsaLIsamrn SsAvqkF
rhoGAP^MRX SQTLQSFQFDFIGDTLTDDEINIAESFKEFAELLNEVENERMMMVHNASDLLIKPLENFR
CELT04C95 AKTLSEFKFETNGTNQTDDERFIATCLKEFGSLINQVEDERMKLVGQAEESYLEPIKKFR
EST483210 SQTLQSFQFDFIGDTLTDDEINIAESFKEFAELLNEVENERMMMVQNASDLLIKPLETFR
EST387042 SQSLQDFQFECIGDAETDDEISIAQSLKEFARLLIAVEEERRRLTQNANDVLIAPLEKFR
consensus sqtLqsFqFdfiGdtlTDDEinIAes KEFaeLlneVEnERmmmvqnAsdllikPlekFR
rhoGAP^MRX KEQIGFT.KERKKKFEKDGERFYSLLDRHLHLSS 153
CELT04C95 TEAIGRTLKEEEKNKKDKESSKFYQTLEKHLHLST 154
EST483210 KEQIGFT.KERKKKFEKDGERFYSLLDRHLHLSS 153
EST387042 KEQIGAA......KLIVYILLFTSYSPDN..... 143
consensus kEqIGft kerkkKfekdgerFys11drh1h1ss

FIG. 3c

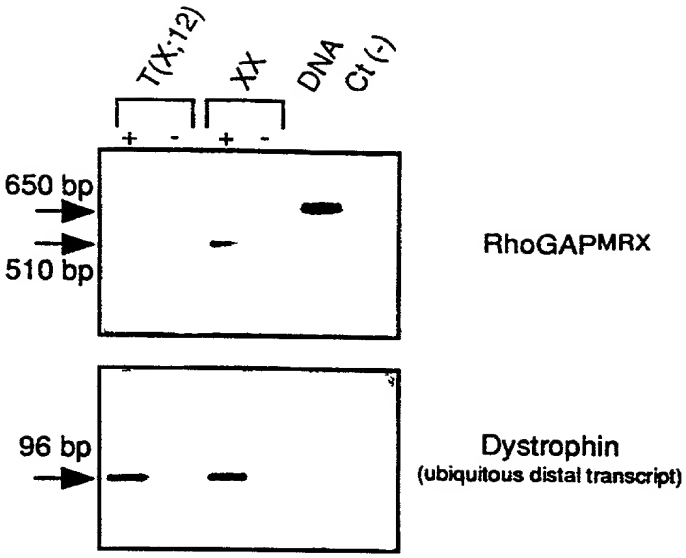
T(X;12)
XX
DNA
Ct(-)
+ - + -
650 bp
510 bp
RhoGAPMRX
96 bp
Dystrophin
(ubiquitous distal transcript)

FIG. 4a

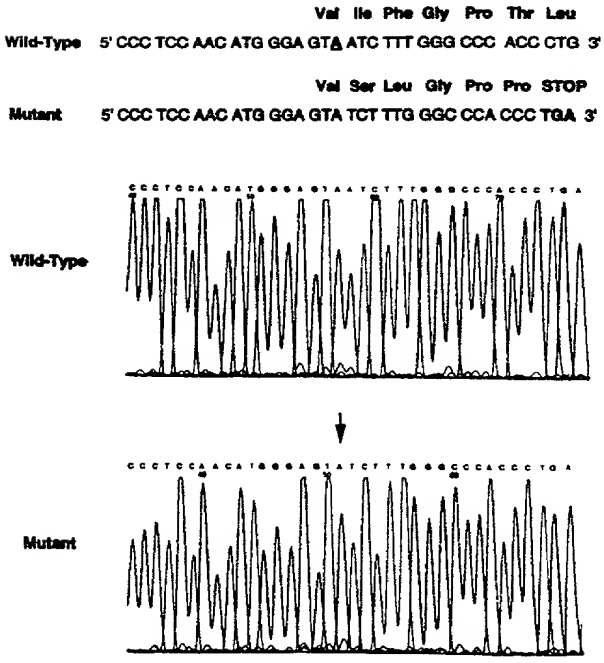
Val Ile Phe Gly Pro Thr Leu
Wild-Type 5' CCC TCC AAC ATG GGA GTA ATC TTT GGG CCC ACC CTG 3'
Val Ser Leu Gly Pro Pro STOP
Mutant 5' CCC TCC AAC ATG GGA GTA TCT TTG GGC CCA CCC TGA 3'
Wild-Type
Mutant

FIG. 4b

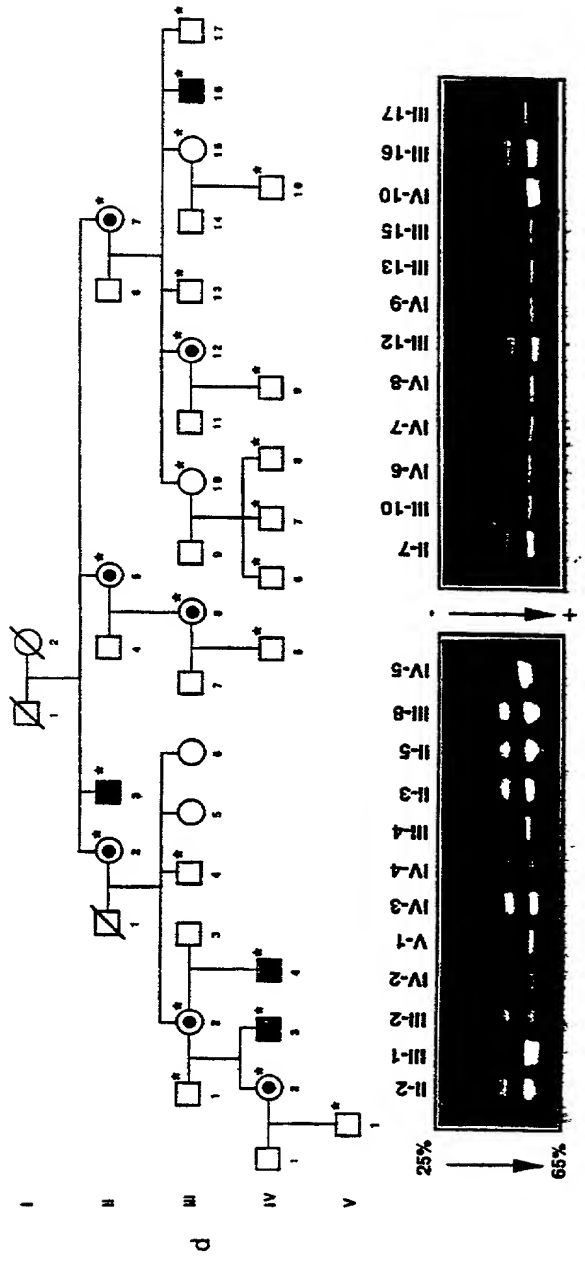
I
II
III
IV
V
II-7
III-10
IV-6
IV-7
IV-8
III-12
IV-9
III-13
III-15
IV-10
III-16
III-17
-
+
II-2
III-1
III-2
IV-2
V-1
IV-3
IV-4
III-4
II-3
II-5
III-8
IV-5
25%
65%

FIG. 4c

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent

In re patent application of: CHELLY et al.

Serial No.: 09/581,422 Examiner:

Filed: June 13, 2000 Art Unit:

For: A NEW GENE CALLED OLIGOPHRENIN 1, ITS... Atty. Docket No.:
P06780US0/BAS

CHANGE OF CORRESPONDENCE ADDRESS CUSTOMER NUMBER DESIGNATION

Honorable Assistant Commissioner for Patents
Washington, D.C.
S I R:

Henceforth, please **change the correspondence address** of the above identified application to the correspondence address associated with the CUSTOMER NUMBER identified below, or to the (same) correspondence address shown below if the Customer Number designation cannot be used.

LARSON & TAYLOR, PLC
Transpotomac Plaza
1199 North Fairfax Street
Suite 900
Alexandria, VA 22314

or

CUSTOMER NUMBER
00881

In addition, please also **appoint the practitioners** (of LARSON & TAYLOR, PLC) associated with this Customer Number to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

If there is any problem in changing the correspondence, please contact the undersigned immediately by telephone at 703-739-4900.

Respectfully submitted,

Date: 21 November 2000

By: B. Aaron Schulman
Registration No.: 31,877

Thesm. 97h02000.1

ATTORNEY/DOCKET NO: _____

DECLARATION FOR PATENT APPLICATION AND APPOINTMENT OF ATTORNEY

As a below named inventor, I hereby declare that my residence, post office address and citizenship are as stated below next to my name; I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention (Design, if applicable) entitled "A new gene called oligophrenin 1, its expression product, and the diagnostic and therapeutic applications thereof".

the specification of which (check one):

☐ is attached hereto.

☒ was filed on June 13, 2000 as Application Serial No. 09/581,422, and was amended on _____ (if applicable).

☐ was filed on 14/12/1998 as International Application (PCT) No. PCT/EP9808557, and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above. I acknowledge the duty to disclose information which is material to the examination of this application in accordance with *Title 37, Code of Federal Regulations, §1.56(a)*. I hereby claim foreign priority benefits under *Title 35, United States Code §119* of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

PRIOR FOREIGN APPLICATION(S)			*PRIORITY CLAIMED*	
97403050.4 (Number)	EUROPE (Country)	December 15, 1997 (Day/Month/Year Filed)	☒ Yes	☐ No
(Number)	(Country)	(Day/Month/Year Filed)	☐ Yes	☐ No
(Number)	(Country)	(Day/Month/Year Filed)	☐ Yes	☐ No
(Number)	(Country)	(Day/Month/Year Filed)	☐ Yes	☐ No

I hereby claim the benefit under *Title 35, United States Code, §120* of any United States application(s) or PCT international application(s) designating The United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of *Title 35, United States Code, §112*, I acknowledge the duty to disclose material information as defined in *Title 37, Code of Federal Regulations, §1.56(a)* which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

Application Number	Filing Date	Status - Patented, Pending or Abandoned

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under *section 1001 of title 18 of the United States Code* and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: I (We) hereby appoint as my (our) attorneys, with full powers of substitution and revocation, to prosecute this application and transact all business in the Patent and Trademark Office connected therewith: Andrew E.Taylor,Reg 7000,Thomas P.Sarro 19396-Harold L.Novick 26011-Mark J.Guttag 33057,Walter C.Gillis 22086,Ross .Hunt Jr 25082,Douglas E.Jackson 28518,Daniel C.Mallery 33532,Marvin Petry 22752,William E.Jacks 4016, B.Aaron Schulman 31877

Send correspondence to: LARSON and TAYLOR - Transpotomac Plaza - 1199 North Fairfax street, suite 900 - Alexandria VIRGINIA 22314 - USA

Telephone Calls to: 703 920 7200.

☐ *See following pages for additional joint inventors.*

Full Name of First or Sole Inventor	*Citizenship*
CHELLY Jamel	French
Residence Address Street	*Post Office Address Street*
11 rue Truillot FRX	
City IVRY SUR SEINE	*City*
State or Country France *Zip* 94200	*State or Country* *Zip*
DATE November 2, 2000	*SIGNATURE*

2-0

Full Name of Joint Inventor	Citizenship
BILLUART Pierre	French
Residence Address Street: 5 rue Poliveau FRX	Post Office Address Street
City: PARIS	City
State or Country: FRANCE Zip: 75005	State or Country Zip
DATE: November 2, 2000	SIGNATURE

3-0

Full Name of Joint Inventor	Citizenship
KAHN Axel	French
Residence Address Street: 10 rue du Docteur Roux FRX	Post Office Address Street
City: PARIS	City
State or Country: FRANCE Zip: 75015	State or Country Zip
DATE: November 2, 2000	SIGNATURE

4-0

Full Name of Joint Inventor	Citizenship
BELDJORD Cherif	French
Residence Address Street: 11 Bis rue Jacob Courant FRX	Post Office Address Street
City: POISSY	City
State or Country: France Zip: 78300	State or Country Zip
DATE: November 2, 2000	SIGNATURE

Full Name of Joint Inventor	Citizenship
Residence Address Street	Post Office Address Street
City	City
State or Country Zip	State or Country Zip
DATE	SIGNATURE

Full Name of Joint Inventor	Citizenship
Residence Address Street	Post Office Address Street
City	City
State or Country Zip	State or Country Zip
DATE	SIGNATURE

☐ See following pages for additional joint inventors